

**Manuscript version: Author's Accepted Manuscript**

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

**Persistent WRAP URL:**

<http://wrap.warwick.ac.uk/128734>

**How to cite:**

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk).

Running head: STRATEGIC NON-DISCLOSURE

Mind The (Information) Gap:

Strategic Non-Disclosure by Marketers and Interventions to Increase Consumer Deliberation

Sunita Sah<sup>1</sup> and Daniel Read<sup>2</sup>

<sup>1</sup>Johnson Graduate School of Management, Cornell University

<sup>2</sup>Warwick Business School

*In Press. Journal of Experimental Psychology: Applied*

Correspondence to:

Dr. Sunita Sah

Johnson Graduate School of Management, Cornell University

Sage Hall

Ithaca, NY 14853

USA

Email: [Sunita.Sah@cornell.edu](mailto:Sunita.Sah@cornell.edu)

Acknowledgments: A six page summary of a prior draft of this manuscript was published in the [Academy of Management Best Paper Proceedings](#) (2017, p 12839), and some of the findings are discussed in a Harvard Business Review article called “[Missing Product Information Doesn’t Bother Consumers as Much as It Should](#)” (September 28, 2017). This manuscript also benefited from helpful discussions with conference attendees at the 2016 Behavioral Decision Research in Management Conference. Many thanks to Dan Feiler, Krishna Savani, and David Butler for feedback on prior drafts of this manuscript.

### Mind The (Information) Gap:

#### Strategic Non-Disclosure by Marketers and Interventions to Increase Consumer Deliberation

#### **Abstract**

Marketers have a choice of what to tell consumers and consumers must consider what they are told or not told. Across six experiments, we show that consumers fail to differentiate between deliberate and non-deliberate missing information (strategic naïveté) and make generous inferences when they do notice missing information is deliberately withheld (charitability). We also show how marketers can take advantage of this by withholding information. We investigate both sides to (1) show the effects of interventions to encourage consumers to consider deliberate nondisclosure in a less naïve and charitable fashion, (2) demonstrate when marketers should disclose (or not) if consumers are naïve and charitable (i.e., breakeven points), and (3) explore the reasons marketers give for (non-) disclosure and consumers' thoughts on why information is missing. Consumers respond differently to distinct but theoretically equivalent framings that increase the salience of non-disclosure. Only when non-disclosure was highly salient, and consumers could compare multiple profiles side-by-side did consumers believe the nondisclosed information to be the worst possible.

**Public significance statement:** Marketers often deliberately withhold relevant information from consumers about the quality of a product or service they are providing. Across six experiments, we show that consumers fail to realize that the information was deliberately withheld, and even when they do realize it, they react negatively only when the missing information is highlighted as deliberately withheld. When disclosure of information is deemed an important policy goal, voluntary disclosure is unlikely to be an adequate substitute for mandatory disclosure.

**Keywords:** *theory of rational disclosure, unraveling of information, voluntary disclosure, decision making, judgment, interventions, choice architecture, equilibrium, strategic behavior, behavioral game theory.*

Imagine a surgeon choosing to display information about herself, including her patient mortality rates, on her online profile. Should she reveal these mortality rates, or should she withhold them? Now, imagine the same scenario from the perspective of the patient who reads the physician's profile, and finds the mortality rate information missing. What will the patient infer about the mortality rates?

The patient's challenge of interpreting missing information is characteristic of many situations in which a consumer must evaluate a situation without some potentially relevant information. The challenge is made greater because the interpretation depends on why the information is absent. There are two broad classes of missing information, which we call "deliberate" or "non-deliberate." Deliberate missing information is when the provider actively withholds information because he does not want you to have it, and non-deliberate is otherwise (perhaps due to an accident, oversight or technical issue). Usually non-deliberate missing information does not justify the same interpretation as deliberate missing information. If you are evaluating a CV and accidentally spill a bottle of ink on it, you should not infer the information concealed by the spillage was particularly bad or good. On the other hand, if the applicant had deliberately spilled the ink just before you read an important passage, you might suspect that the applicant did *not want you to see* the missing information and you can therefore infer the information is probably bad.

In this paper, we focus on cases of *deliberate* non-disclosure, in which information providers or marketers intentionally fail to disclose information that they possess and might be relevant to their consumers. Such situations are commonplace. For example, when buying a car, going to a restaurant, deciding on a movie, or finding a doctor, you will be presented with selected information about product features or service characteristics. In these situations,

we investigate whether consumers *notice* when information is withheld, and the *inferences* they make when they do notice, or when their attention is drawn to withheld information.

We document the effects of recipients not differentiating between deliberate and non-deliberate missing information (an effect we call “strategic naïveté”) and of making generous inferences when they do notice missing information is deliberately withheld (an effect we call “charitability”). We investigate the strategic behavior of both information providers and recipients and (1) demonstrate interventions that can help recipients notice and make more skeptical inferences about deliberate non-disclosure, (2) identify when providers should withhold information to obtain the most favorable response from recipients (i.e., the “breakeven point” for providers to disclose), and (3) describe the reasons providers give for (non) disclosure and recipients give for why they believe information is missing. The interventions we examine include different framings of deliberate non-disclosure, perspective-taking, and comparative processing. These interventions moderate recipients’ reactions to deliberate non-disclosure.

### **Interpretation of Missing Information**

Our focus is on two important research problems. First, how do providers decide whether to disclose what they know? Second, how do recipients respond to deliberate non-disclosure?

The decision of providers to disclose and the decision by recipients of how to respond are made sequentially, and both depend on the other. The information provider decides first whether to disclose, but their decision will (or should) depend on what they anticipate the recipient will do. For example, a marketer chooses how to describe a product before consumers decide whether to buy it. But the marketer chooses their description based on how they expect consumers will respond. In this sense, therefore, the anticipated response of the recipient is logically prior to the actions of the provider. For example, imagine a student who

has deliberately chosen to disclose three grades out of a possible four, each of which are on a five-point ABCDF scale. She reports the set  $\{A, A, B\}$ , leaving another grade  $X$  undisclosed. How will a recipient (perhaps a potential employer or professor) evaluate this student?

There are numerous possible responses. One extreme response is the maximum skepticism predicted by a normative game theory analysis that assumes “strategic rationality” – that is, unlimited and perfect strategic reasoning on both sides, including perfect strategic reasoning about the perfect strategic reasoning of the other side (Mariotti, 1995; Selten, 1988). Game theory predicts a process often referred to as information *unraveling*, which is that the recipient will infer that the provider’s performance on any undisclosed attributes is the worst possible (Grossman, 1981; Grossman & Hart, 1983; Milgrom, 1981; Peppet, 2011; Viscusi, 1978). This means the student’s true profile will be estimated as  $\{A, A, B, F\}$ . This occurs because all providers except those with the worst performance will want to avoid being pooled (or “confused”) with those having poorer performance than them. In other words, if the student’s last grade was an A, she presumably would have disclosed it to avoid recipients believing she may not have an A. What about a B? If the student was sure to disclose an A, then a B is the best of all the grades that remain undisclosed, so the argument to disclose an A also leads her to disclose a B. By the same argument, if she has a C (or even a D) grade she will also disclose because, if she doesn’t, the recipient will infer she has the worst possible grade of F. The iterative nature of this process is known as the “unraveling” of information.

The prediction of unraveling and full disclosure is built on several assumptions: (a) the provider possesses the information and cannot credibly misrepresent it; (b) disclosure is costless for the provider; (c) on each dimension, ranks can be defined unambiguously so that all recipients agree what constitutes high and low quality; (d) the range of possible quality values is known to both the provider and recipients, and (e) conditions (a) through (d) are

common knowledge. These conditions are likely to prevail in many familiar contexts, but not all.

Thus, by changing some assumptions, the game theory analysis can allow more moderate predictions than  $\{A, A, B, F\}$ . For instance, a student might not provide complete transcripts because they are costly to obtain (a violation of  $b$ ). If a potential employer is aware of this cost, they might attribute absent transcripts to an inability to bear this cost. Non-disclosure will also not be judged as harshly if (for example) grades of C, D and F are considered equally bad (a violation of  $c$ ), since then the non-disclosed grade could be as high as a C (for more on this topic, see Dranove & Jin, 2010, and references therein).

Another extreme response is the maximally myopic response which effectively ignores any missing information (Simmons & Lynch, 1991), and so treats the provider as someone who has provided all the relevant facts. The recipient would therefore treat the student as if she had taken only three courses for which she received  $\{A, A, B\}$ . This response captures the essence of Kahneman's (2011) phrase, "what you see is all there is."

Intermediate cases occur when the relevance of missing information is recognized, but the missing information is treated as if it were *not deliberately withheld*. One possibility is that the recipient estimates what a typical student who has two As and a B would likely receive as a fourth grade, perhaps considering an average of these grades or using the most frequent observation as the most likely missing grade (c.f., Garcia-Retamero & Rieskamp, 2009). Another intermediate response is to treat the three grades as a sample from a population who all report their three highest grades (limited information may be deliberate to not overwhelm recipients with too much information). The missing grade would then be no higher than the lowest of the three reported, with the specific prediction depending on the recipient's understanding of how grades are distributed.



Of course, these are not the only ways a recipient might respond to undisclosed information, but the range of easily-imaginable responses demonstrates that even a highly sophisticated information provider is faced with a complex challenge. The success of her disclosure policy will depend on how well she predicts the response of the recipient to her policy. A sophisticated provider will choose a disclosure strategy to induce the recipient to evaluate her in the best possible light.

### **Empirical Evidence**

Although a variety of responses to missing information have been documented and explored empirically and analytically (Garcia-Retamero & Rieskamp, 2009), there is as yet no systematic investigation of how recipients and providers make decisions concerning *deliberate* non-disclosure. Some experimental studies have either focused on *non-deliberate* non-disclosure, or else subsumed both deliberate and non-deliberate missing information into a single category (Hernandez, Han, & Kardes, 2014; Kardes, Posavac, & Cronley, 2004) and looked at the inferences consumers make from such missing information (Johnson & Levin, 1985; Ross Jr & Creyer, 1992; Sanbonmatsu, Kardes, & Herr, 1992; Sanbonmatsu, Kardes, Houghton, Ho, & Posavac, 2003; Sanbonmatsu, Kardes, Posavac, & Houghton, 1997; Sanbonmatsu, Kardes, & Sansone, 1991; Simmons & Lynch, 1991). Our focus in this paper is on decisions made by recipients and providers when information is *deliberately withheld*. We contribute to this literature while focusing on the interlinked interpretation and decisions of recipients and providers. We explore differing interventions to make recipients more strategic and we also document providers' reasons for disclosing (or not disclosing) as well as recipients' thoughts regarding the missing information.

In the "outside world," missing information is commonly withheld deliberately. Much of the field literature reveals that providers given a chance do not usually disclose to the level predicted by skeptical game theorists. For example, Mathios (2000) studied whether salad

dressings disclosed their fat content and found that while low fat dressings (under 6 grams per serving) almost always disclosed, only 60% of those with intermediate fat content, and 20% of those with high fat content did. Similarly, Carrillo, Cellini, and Green (2013) found that homeowners are highly selective about reporting the quality of local schools, and that disclosure increases with school quality. And Butler and Read (2017) show that hotels are less likely to disclose their TripAdvisor ratings if the ratings are low.

Also contrary to the strong game theory predictions, when disclosure was voluntary, marketers responded in ways suggesting they do not believe consumers are fully strategic. For example, Jin and Leslie (2003) studied the effect of a new hygiene rating scheme for restaurants. Because of varying legislation, displaying the rating was obligatory in one county but voluntary in another. Restaurants in both counties improved their hygiene relative to the pre-rating era, but those who were obliged to disclose improved more. Those not obliged acted as if they expected consumers to not engage in the full unraveling reasoning process or else to be “charitable” about withheld information. It is noteworthy that an obligation to disclose led to improved performance, a finding that has recently been replicated in the laboratory for both mandatory and voluntary disclosure (Sah & Loewenstein, 2014).

Brown, Camerer, and Lovallo (2012, 2013) describe a similar case. Film studios often release movies they expect to receive poor reviews on without offering reviewers a pre-release screening (that is, without disclosing the movie’s quality). Moviegoers who apparently give insufficient weight to the implications of non-disclosure initially attend these movies in higher numbers than they would have if the movies had actually received a poor review. A lack of full unraveling has also been documented in other domains, including health maintenance organizations (HMOs) (Jin, 2005) and business-school rankings (Luca & Smith, 2015). These and other field studies (Fung, Graham, & Weil, 2007) show that people do not always infer the worst and tend to be strategically naïve and/or charitable.

Under some circumstances, recipients do act more like maximally skeptical game theorists. Forsythe, Isaac, & Palfrey (1989) conducted a multi-round asset-bidding experiment in which providers could choose to reveal the quality of the product they were selling. Consumers did not initially assume the worst from undisclosed information, but after several rounds they did. By contrast, Jin, Luca, and Martin (2015) conducted a series of multi-round experiments and did not find evidence of full unraveling. John, Barasz, and Norton (2016) found that, in a laboratory setting, people did assume the worst about individuals who selected “choose not to answer” when asked to disclose personally sensitive information (e.g., whether they had ever cheated on a tax return, the lowest grade they had ever received). Thus, a range of responses to undisclosed information has been documented, from the extremes of complete unraveling observed by John et al. (2016) to the very low level of unraveling described by Brown et al. (2012, 2013).

### **Strategic Naiveté and Charitability**

The distinction between deliberate and non-deliberate missing information is sometimes subtle, but always important. As mentioned, prior research has documented some insensitivity to non-deliberate missing information. That is, when information is removed randomly or there is no reason to assume it was withheld for reasons of concealment. For example, papers on omission neglect sometimes subsume both deliberate and non-deliberate non-disclosure into a single category (Hernandez et al., 2014; Kardes et al., 2004). Our work is relevant to omission neglect, but we focus on *deliberately* withheld information. Only deliberate non-disclosure requires strategic analysis from the recipient. Therefore, we focus on contexts in which the provider wants to sell something to the recipient, such as a product or service. In these situations, providers are likely to have deliberated on what information to withhold or not.

We predict, however, that recipients will often not differentiate between the deliberate and non-deliberate absence of information, and providers will often not consider the responses of recipients to missing information when deciding whether to disclose or not. Our analysis is related to the concept of “Level-k” thinking (e.g., Camerer & Johnson, 2004; Nagel, 1995; Stahl II & Wilson, 1994), a key concept in modern game theory. The core idea is that strategic actors form a heterogeneous population that differ in their depth of strategic reasoning. Some are entirely naïve, displaying what is often called “Level-0” reasoning, and do not consider the decisions of others when making strategic choices. For instance, in our setting, recipients might fail to consider why information is missing, or providers might fail to consider the recipient’s response to missing information. Others may consider the reasons for the nondisclosure decisions of others, but they do not complete the unraveling process or they conduct relatively innocuous attributions. For example, Benndorf, Kübler & Normann (2017) show, in a labor market context in which workers can reveal their level of productivity, that very few workers and employers undertake complete unraveling. As Benndorf (2018, p. 23) puts it, “Workers are generally reluctant to disclose their private information, and employers enforce this behavior by bidding less competitively if workers reveal compared to the case where they conceal information.”

The problem of interpreting deliberate non-disclosure is strategically complex. First recipients must recognize information is missing. Next, recipients must both recognize that missing information has been *deliberately* withheld and then infer what the missing information is likely to be based on considerations of why that information was withheld. We anticipated that unless the reason for information non-disclosure is made salient, recipients would fail to infer that information was absent because the provider deliberately withheld it.

H1a: Strategic naiveté: Consumers will not differentiate between deliberate and non-deliberate non-disclosure unless missing information is made salient.

Furthermore, when the deliberateness of non-disclosure is made salient, we anticipate that merely knowing that information was deliberately withheld is insufficient for recipients to infer the worst about it. Rather, they will display something analogous to “Level-1” reasoning (Camerer & Johnson, 2004), in which they respond partially to the missing information, but will not conduct the full strategic game theoretic analysis.

H1b: Charitability: Consumers will be charitable towards deliberately missing information.

It should be noted that if providers are frugal about providing information then consumers will not necessarily be wrong in their inferences that a non-disclosing provider is not the worst possible. We return to this point in Experiments 5 and 6 and in the General Discussion.

### **Interventions to Increase Consumer Deliberation**

Given that consumers are likely to be strategically naive and charitable towards deliberate non-disclosure, can we encourage more strategic thinking? We investigate different interventions aimed at encouraging consumers to make more strategic inferences about missing information. We show the effect of several choice architecture interventions that attempt to make deliberate non-disclosure more salient. These involve different framings of non-disclosure, perspective-taking interventions that prompt consumers to deliberate more about the reasons for missing information, and comparative processing strategies that enable consumers to compare information from multiple providers, only some of whom disclose. We expected that although some of these interventions would encourage deeper strategic thinking, consumers would never reach full unraveling and would therefore not overturn the predictions of H1b. However, if most people think only one step ahead (as suggested by Camerer & Johnson, 2004), then an extra step of deduction may improve decision-making.

H2: Interventions such as different framings of deliberate non-disclosure, perspective-taking and comparative processing will moderate consumers' reactions to missing information and lead to partial unraveling.

### **Providers' Disclosure Decisions**

We propose that non-strategic thinking will not only be confined to consumers but will also be found in the thinking of providers. We predicted that providers would typically *not* consider the possibility of strategic thinking by consumers. Providers would therefore base their disclosure decisions on the principle of withholding information they don't want consumers to know, rather than on the inferences consumers could make if they withhold information.

H3: Strategic naiveté: Providers will not base their disclosure decisions on what consumers would infer from the *absence* of deliberately withheld information.

We therefore anticipated “strategic failures” on the part of both providers and consumers in response to deliberate non-disclosure: providers would disclose less than they should according to strategic rationality and consumers would treat non-disclosures less skeptically than game theory would predict. To test these predictions, we examine both providers and consumers and estimate a “breakeven point” above which it is beneficial for providers to disclose, and below which it is beneficial for them to conceal information. We find that contrary to the strong predictions of game theory, providers should withhold quite a bit of information and that a “Level-0” provider is likely to be much better off than a “Level-0” consumer.

### **Overview of Experiments**

We conducted six experiments exploring decisions by recipients (Experiments 1 to 4) and providers (Experiments 5 and 6). We embedded the problem in a medical context. This

domain provides an apt context to test our hypotheses for several reasons. The primary reason is that the physician-patient domain is a high-stakes setting that is familiar and accessible to almost everyone. In addition, online physician profiles are increasing in use all over the world, both for advertising and quality control even if little is known about how consumers assess them (Hanauer, Zheng, Singer, Gebremariam, & Davis, 2014). There is also an ongoing debate concerning whether physicians should have the option of withholding information and, if so, whether they are likely to do so if they believe it is in their best interest (Gray, 2013; Jones, Gabe, & Exworthy, 2013). Finally, debates over physician trustworthiness are becoming increasingly important to physicians, regulators, and patients (Grande, Shea, & Armstrong, 2012; Rose et al., 2019; Sah, 2012; Thom, Kravitz, Bell, Krupat, & Azari, 2002).

In Experiment 1, we examine whether participants distinguish between non-deliberate non-disclosure (when information is randomly withheld) and deliberate non-disclosure (when information is actively withheld). We also conduct our first test of a different framing for deliberate non-disclosure to increase the salience of non-disclosure and decrease the strategic naiveté and charitability of consumers. In Experiment 2, we use a greater range of framings of deliberate non-disclosure as well as asking participants to explicitly reflect on the missing information. Experiment 2 also allows us to estimate a breakeven point above which it is beneficial for providers to disclose, and below which it is beneficial for them to conceal information. Experiment 3 examines whether perspective-taking (i.e., consumers considering why providers did not disclose information) decreases naiveté and charitability. Experiment 4 tests consumers' reactions to salient non-disclosure and the comparison of multiple profiles. Experiments 5 and 6 focus on the provider, asking what decisions the provider will make and searching for evidence of strategic thinking to better understand the reasoning behind deliberate non-disclosures.

In all experiments we report results from all participants recruited, experimental conditions, and measures. We based our sample size decisions on personal views about “adequate” sample sizes (100 or more), opportunity and budget. All analyses were conducted only after data collection for that experiment was completed. Data and materials are available at <https://osf.io/g9hpf>. Our study received ethics (Institutional review board) committee approval.

### **Experiment 1: Comparing Deliberate and Non-Deliberate Missing Information**

In this experiment, we investigated whether recipients would distinguish between randomly missing and deliberately not-provided information. We predicted they would not (strategic naiveté, H1a). We also tested whether increasing the salience of non-disclosure deliberateness would increase its impact (H2). We predicted that even with more salient non-disclosure, consumers would interpret withheld information in a more positive light than predicted by unraveling theory (charitability, H1b).



## Method

**Participants and design.** Participants from Mechanical Turk<sup>1</sup> ( $N = 493$ ; 173 women, 320 men;  $M_{age} = 29.7$ ,  $SD = 9.2$ ) were randomly assigned into one of three conditions: *Random*, *Not Provided*, and *Refused*.<sup>2</sup>

**Procedure.** Participants role-played patients searching for a doctor online (see Figure 1). They were informed that all doctors in a major U.S. city received patient ratings from 0 (extremely poor) to 100 (outstanding) on the five dimensions of quality of care, trustworthiness, availability, bedside manner, and value for money. For each dimension, we displayed the average combined ratings for all doctors in the city. These *average ratings* ranged from 78 to 85. Participants were next informed they would see only four of the five ratings received by a certain Dr. Y. We varied which dimension was missing and how its absence was explained. In the *Random* condition participants were informed that “we randomly removed” the rating; in the *Not Provided* condition they were informed that “Dr. Y did not provide” the rating; and in the *Refused* condition that “Dr. Y refused to provide” it. Note that the *Refused* condition was a more salient version of the *Not Provided* condition, but both conditions represent deliberate non-disclosure.

---

<sup>1</sup> Amazon Mechanical Turk (MTurk) is an online labor system that allows researchers to conduct online experiments with a sample more diverse than undergraduate students; (see Goodman, Cryder, & Cheema, 2012; Rand, 2012). In all experiments using MTurk in this paper, we included attention filters that participants had to pass before being allowed to start the study as well as a question of commitment “Sometimes survey participants skip the instructions and rush through, which ruins our results. Are you willing to actually read the instructions on the next page?” Participants had to confirm “Yes” before starting the study. We also included at least one attention filter during the study itself which all participants passed unless noted.

<sup>2</sup> Six participants had response time durations 3 standard deviations (SD) above the mean. Removing those participants do not change the pattern or significance of the results.

Participants estimated the missing rating and then answered questions about Dr Y. First, “Do you think Dr. Y is likely to be better or worse than the average doctor on the missing dimension?” answered on a 7-point scale from very much worse (1) to very much better (7). Next, they assessed Dr. Y’s competence, credibility and integrity on separate 7-point scales. We averaged these three ratings to give an overall measure of Dr. Y’s quality (Cronbach’s  $\alpha = .90$ ).

## Results

**Estimate of the missing rating.** Estimates varied significantly among the three conditions,  $F(2, 490) = 38.22, p < .001, \eta_p^2 = .14$  (Figure 2). Supporting H1a, in the *Random* condition, the mean estimate was 77.02 ( $SD = 5.26$ ), which was not significantly higher than that in the *Not Provided* condition, ( $M = 75.94; SD = 6.06$ ),  $t(490) = 1.31, p = .19$ . Displaying strategic naiveté, consumers did not distinguish between non-deliberate (random) and deliberate non-disclosure (H1a). The more salient non-disclosure in the *Refused* condition did, however, produce significantly lower estimates ( $M = 70.36; SD = 10.02$ ) than the *Not Provided* condition,  $t(490) = 6.80, p < .001$ , despite both framings representing deliberate non-disclosure.

The standard deviation for the estimate of the missing dimension in the *Refused* condition (10.20) was significantly larger than in the *Not Provided* (6.06) and the *Random* conditions (5.26); Levene’s test of equality of error variances,  $F(2, 490) = 20.34, p < .001$ . However, the violation of the assumption needed to conduct analysis of variance was not great (the standard deviations differed by less than double the smallest standard deviation) and thus statisticians typically recommend univariate analysis will still be robust and reliable (Glass, Peckham, & Sanders, 1972). However, we also conducted non-parametric tests (the

Independent-Samples Kruskal-Wallis Test), which revealed a similar pattern of results.<sup>3</sup> A larger variance in the *Refused* condition is consistent with the possibility that responses in this condition represents a mixture of people who ignore the implications of deliberate non-disclosure and those who respond to it.

For the five dimensions, all but one (availability) displayed the same significant pattern as reported for the aggregate of the missing dimensions (eFigure 1), all  $ps \leq .001$ . The availability dimension did not differ significantly between conditions,  $F(2, 95) = 1.18, p = .31, \eta_p^2 = .02$ . The largest difference between the *Refused* condition and the other two conditions was seen for the trustworthiness dimension,  $F(2, 96) = 16.83, p < .001, \eta_p^2 = .26$ . Some dimensions are therefore more heavily weighted if missing (e.g., trustworthiness) than others (e.g., availability). This is interesting, since one critical assumption of the unraveling prediction is that all consumers agree what constitutes high and low quality. While trustworthiness is a distinct unambiguous concept to most patients, the physician's availability may be a dimension that does not mean the same thing to everybody. That is, while everyone will generally want the physician to be available to them personally, they may also want the physician to be in demand.<sup>4</sup>

**Dr. Y: Better or worse than average?** Judgments of whether Dr. Y was better or worse than the average doctor on the missing dimension showed the same pattern as estimates of the missing dimension. In fact, these judgements were significantly correlated

---

<sup>3</sup> Similarly, for other experiments reported in this manuscript, the Independent-Samples Kruskal-Wallis Test was conducted whenever there was a violation of equal variances. Again, the violations were not strong and similar patterns of results were obtained with the non-parametric analyses.

<sup>4</sup> For example, imagine a nightclub that has to disclose the average length of the queue. For some people a short queue would be desirable since it will let them get in easily; for others a long queue is better because it means the nightclub is popular.

with how far the participants' estimate of the missing dimension was from the average rating of all doctors,  $r = .48, p < .001$ .

The three conditions differed significantly,  $F(2, 490) = 15.51, p < .001, \eta_p^2 = .06$ . Ratings in the *Random* and *Not Provided* conditions did not differ (respectively,  $M = 3.84, SD = .85$ ;  $M = 3.70, SD = .87$ ;  $t(490) = 1.38, p = .17$ ), again providing support for the hypothesis that consumers do not distinguish between deliberate and non-deliberate nondisclosure (H1a). However, different framings of deliberate non-disclosure had an effect (H2): ratings of the doctor were substantially lower in the *Refused* condition ( $M = 3.31, SD = 1.02$ ) than in the *Not Provided* condition,  $t(490) = 3.96, p < .001$ .

Further, in both the *Not Provided* and *Refused* conditions, consumers were charitable (H1b)—i.e., they did not rate Dr. Y to be very much worse than the average doctor (1 on the 7-point scale) on the missing dimension (for *Not Provided*,  $t(161) = 39.57, p < .001$ ; for *Refused*,  $t(169) = 29.62, p < .001$ ). However, the mean in all conditions was below the mid-point of the scale (4), all  $ps < .02$ , suggesting that participants rated Dr. Y worse than the average doctor on the missing dimension.

**Dr. Y's quality.** A similar pattern was seen with the assessments of competence, credibility and integrity (averaged) across the three conditions,  $F(2, 490) = 20.79, p < .001, \eta_p^2 = .08$ . The *Random* and *Not Provided* conditions did not differ (respectively,  $M = 5.10, SD = .95$ ;  $M = 5.11, SD = .95$ ;  $t(490) = .03, p = .98$ ), providing further support for H1a. However, the *Refused* condition was significantly lower ( $M = 4.48, SD = 1.17$ ) than the *Not Provided* condition,  $t(490) = 5.58, p < .001$  (H2). In both deliberate non-disclosure conditions, consumers were charitable (H1b) and did not assume the worst about Doctor Y's quality, (for *Not Provided*,  $t(161) = 55.03, p < .001$ ; for *Refused*,  $t(169) = 38.82, p < .001$ ).

## Discussion

As predicted by H1a, consumers displayed strategic naïveté in that evaluations did not differ between information absence due to randomness (*Random* condition) and deliberate non-disclosure (*Not Provided* condition). Only when deliberate non-disclosure was made salient (*Refused* condition), did consumers attend to it and give lower ratings to the deliberately non-disclosing doctor (H2). Even in this condition there was charitability, in that consumers did not assume the worst about the doctor, even in the *Refused* condition (H1b).

## Experiment 2: Breakeven Point for Non-Disclosure: Strategic Naïveté and Charitability

In this experiment, all non-disclosures were deliberate. As well as investigating strategic naïveté and charitability, we also explore in greater detail interventions to increase attention to deliberate non-disclosure (H2). We tested different framings of deliberate non-disclosure as well as an intervention designed to direct consumers' attention to the undisclosed information by asking them to estimate the missing information. The latter intervention consisted of respondents in all non-disclosure conditions estimating the missing information either before or after stating how likely they would be to choose the provider. We expected that making an estimate of the missing information before stating this likelihood would increase deliberation on why the information was missing, leading to a decrease in strategic naïveté and/or charitability. Consumers would therefore respond more negatively to withheld information and be less likely to choose the provider when they estimated the missing information first. Kardes, et al., (2006) used a similar technique to heighten consumers' cognizance of missing information, and found consumers gave more moderate evaluations when estimating missing information prior to giving an overall judgment.

Finally, we measured what we call “the breakeven point”—the point at which a provider should be indifferent between disclosing and not disclosing information about

themselves, given consumers' responses to nondisclosure. The breakeven point provides us with a quantitative estimate of the effects of strategic naiveté and charitability.

## Method

**Participants and design.** Participants ( $N = 1,104$  MTurkers; 442 women, 638 men, 24 genders unreported;  $M_{age} = 31.3$ ,  $SD = 10.4$ )<sup>5</sup> were randomly assigned to one of nine conditions. One condition was Full Disclosure; the other eight conditions were formed from a four (framing of deliberate non-disclosure) by two (timing of trustworthiness judgment) design. The four levels of framing were *Absent*, *Blank*, *Not Provided*, and *Refused*. Trustworthiness judgments were made either before or after participants estimated their likelihood of choosing the provider.

**Procedure.** Participants role-played patients searching online for a physician. They viewed a profile of Dr. Green (see eFigure 2 in the supplement) and read: "All doctors in your region obtain ratings on five dimensions of medical care from their patients. These ratings are given using a scale from 0–100." For each dimension, participants were told the range of evaluations given to doctors in their region. For the key trustworthiness dimension, this range was 51 to 99; thus 51 was the lowest possible rating. After viewing the profile, participants saw Dr. Green's ratings and were informed that Dr. Green had provided the ratings. Those randomized to the Full Disclosure condition saw one of three trustworthiness ratings for Dr. Green: high (99, the top rating), intermediate (75), or low (52).

In the remaining conditions, the trustworthiness rating was absent and both non-disclosure framing and the sequence in which a trustworthiness rating was made were manipulated. The least salient *Absent* framing presented neither the information nor even the

---

<sup>5</sup> Five participants had response time durations 3 SD above the mean. One participant reported an age of 17. Removing those participants do not change the pattern or significance of the results.

label Trustworthiness; the entire row was absent, so consumers viewed only the four dimensions and ratings. Although consumers were informed that all doctors obtained ratings on five dimensions of medical care, they would not know that the missing rating was trustworthiness nor necessarily even notice that a rating was missing. The *Absent* framing reflects a common version of deliberate non-disclosure in which information is absent without comment and the consumer may not even know it is missing.

With the *Blank* framing, consumers saw a blank box where the missing trustworthiness rating should have been, along with the label Trustworthiness and its range (51–99) (see eFigure 2). Salience of non-disclosure increased with the *Not Provided* and *Refused* framings (parallel to conditions in Experiment 1), where the formerly blank space contained, respectively, the words “Doctor did not provide” and “Doctor refused to provide.”

After reading the profile, participants estimated the likelihood they would choose that doctor on a scale from 0 (definitely not) to 100 (definitely). Participants also provided written explanations of this choice likelihood (see supplement for coding and analyses of these explanations). In each non-disclosure condition, half the participants estimated the undisclosed trustworthiness rating (on a scale from 51 to 99) before providing their choice likelihoods; and half did so afterward. We expected that choice likelihoods would be lower for those who made the trustworthiness estimates first than those who made them second. This applies, perhaps even more so, to the *Absent* framing: consumers cannot be expected to discount when there is no possibility of knowing what information is withheld (*Absent, trustworthiness rating second* condition). In the *Absent, trustworthiness rating first* condition, responding to the trustworthiness estimate rating allows participants to learn that the trustworthiness information was withheld before they made their choice likelihood ratings.

## Results

**Choice likelihood.** Figure 3 shows choice likelihoods for all conditions including the three subgroups of the Full Disclosure condition (High, Intermediate and Low). Reassuringly, the higher the trustworthiness rating, the more likely consumers were to choose Dr. Green. We confirmed this observation with an overall ANOVA on the three subgroups in the Full Disclosure condition,  $F(2, 123) = 54.13, p < .001, \eta_p^2 = .47$ , revealing that the trustworthiness dimension was extremely important for *informed* consumers.

We next examined the non-disclosure conditions. A 4 (Non-disclosure Framing) by 2 (Timing of trustworthiness estimate) ANOVA revealed a main effect of Framing,  $F(3, 970) = 37.16, p < .001, \eta_p^2 = .10$ ; the more salient the framing of non-disclosure, the less likely consumers were to choose Dr. Green. There was no significant effect of Timing,  $F(1, 970) = 3.31, p = .07, \eta_p^2 = .003$ , nor any interaction,  $F(1, 970) = 1.14, p = .33, \eta_p^2 = .004$ . Figure 3 shows the main pattern clearly. We collapsed data over the Timing condition and made pairwise comparisons between adjacent Framing levels. The *Absent* and *Blank* framings were indistinguishable ( $p = .17$ ), demonstrating strong effects of strategic naiveté. The choice likelihoods in both these conditions closely resembled (and were not significantly different from) those of the intermediate trustworthiness subgroup in the Full Disclosure condition. This suggests that when the trustworthiness dimension is missing but not made salient, the doctor is treated as approximately average on that dimension.

The choice likelihoods in both *Not Provided* and *Refused* framings were significantly lower than those in the *Absent* and *Blank* framings; *Not Provided* was significantly lower than *Blank* ( $p = .005$ ) and *Refused* was in turn significantly lower than *Not Provided* ( $p < .001$ ). An important comparison is between the choice likelihood under Full Disclosure when the Dr. Green had low trustworthiness of 52, and the choice likelihood in the highly salient *Refused* condition. In the Full Disclosure condition, the mean choice likelihood was 34.55



( $SD = 25.38$ ). In the *Refused* condition it was 49.42 ( $SD = 25.71$ ). In this environment, therefore, a provider with a very poor trustworthiness rating would be better off withholding that rating (even with the flag that they “refused” to provide it) than disclosing it.

**Trustworthiness estimates.** We conducted a 4 (Non-disclosure Framing) by 2 (Timing of trustworthiness estimate) ANOVA. The more salient the non-disclosure framing, the lower the trustworthiness estimate,  $F(3, 967) = 106.56, p < .001, \eta_p^2 = .25$ : Each increase in framing salience significantly reduced the trustworthiness estimate (all  $p$ 's  $< .001$ ). That is, trustworthiness estimates with the *Absent* framing were higher than with *Blank*, *Blank* was higher than *Not Provided*, and *Not Provided* was higher than *Refused*. Although *Refused* produced the lowest trustworthiness estimate ( $M = 67.27, SD = 9.72$ ) it was still significantly higher than the worst possible trustworthiness score (51),  $t(242) = 26.09, p < .001$ . As predicted, in none of the different non-disclosure framings did consumers estimate that Dr. Green would have the worst possible trustworthiness rating (H1b).

A significant but small effect of Timing was present,  $F(1, 967) = 5.93, p = .02, \eta_p^2 = .006$ . Trustworthiness estimates were approximately just over one point higher when provided before making a choice likelihood estimate ( $M = 74.81, SD = 8.32$ ) than when provided afterwards ( $M = 73.00, SD = 9.58$ ). There was no interaction between the Non-disclosure Framing and Timing,  $F(3, 967) = .23, p = .88, \eta_p^2 = .001$ .

**Calculating the breakeven point of disclosing trustworthiness.** Focusing on Full Disclosure alone, we conducted a linear regression analysis predicting the likelihood of choosing Dr. Green from the disclosed trustworthiness level. This analysis produced the following (significant,  $F(1, 124) = 85.60, p < .001, R^2 = .41$ ) equation:

$$\text{Choice likelihood} = -4.55 + 0.84 * \text{Trustworthiness level}$$

To calculate the breakeven point for disclosing trustworthiness, we took the mean choice likelihood for each level of non-disclosure framing and used the preceding equation to

infer the corresponding level of trustworthiness. The result is the breakeven point, at which a physician with this level of trustworthiness is equally likely to be chosen by disclosing or not disclosing the trustworthiness rating.

For the *Absent* framing, the level of trustworthiness at which a doctor would be equally likely to be chosen whether disclosing or not is 86.60. For the *Blank* framing, it is 83.44; for *Not Provided* it is 76.96; for *Refused* it is 64.25. Above these levels, doctors are more likely to be chosen if they disclose; below them, they are less likely to be chosen.

## Discussion

This study quantifies the naiveté and charitability that people display under different disclosure regimes. Most participants rated the provider close to average or above average for the missing dimension (H1a and H1b). Our salience intervention led to partial unraveling (H2): consumers rated the provider worse when non-disclosure framing was more salient. With the *Absent* framing, where non-disclosure was entirely shrouded, a non-disclosed trustworthiness rating was equivalent to trustworthiness of 87, demonstrating strategic naiveté; this rating fell to 64 when non-disclosure was maximally salient, but was still higher than the worst possible rating (51), demonstrating a high degree of charity. Notably, however, in contrast to Kardes, et al., (2006), asking people to provide a trustworthiness rating before assessing their likelihood of choosing Dr. Green did not reduce that likelihood. This pattern shows that in many circumstances, concealing information will often be the best strategy for providers even when the information they possess is not the worst possible and even when consumers have an opportunity to reflect on why the information is undisclosed.

## Experiment 3: Perspective-taking Increases Strategic Thinking in Consumers

We next tested a further intervention to increase strategic thinking and decrease consumer naiveté and charitability (H2). Recipients deliberated on *why* non-disclosed

information was withheld (taking the perspective of the provider), before (vs. after) giving the likelihood that they would choose the provider and estimating the provider's rating. Since strategic thinking is based on deliberation, we predicted that participants who first considered the responses of providers would more readily reason backwards and be less charitable towards the provider.

In addition to being a manipulation of deliberation, the reasons given by consumers are informative in themselves. Recall that the unraveling prediction is premised on consumers inferring that the provider's decision to withhold information is itself based on beliefs about how the consumer will respond to the absence of information. The reasons we elicit are a snapshot of these beliefs – not the ideal beliefs of the strategically rational consumer of game theory, but the beliefs of actual information recipients considering the absent information. Asking participants to consider and give a reason as to why information may be missing is in itself an intervention that should increase strategic thinking (H2). However, we still anticipated that consumers would show some levels of charitability in their stated beliefs in that they would not state that the provider had the worst possible information (H1b).

## Method

**Participants and design.** Participants ( $N = 709$  MTurkers; 323 women, 383 men, 2 genders reported as “other”, 1 gender unreported;  $M_{age} = 34.8$ ,  $SD = 10.5$ )<sup>6</sup> were randomly assigned to a 2 (Intervention: *Control* vs *Deliberation*) x 2 (Non-disclosure Framing: *Not Provided* vs *Declined*) between-subject design. We expected an interaction such that *Deliberation* would have a greater effect on reducing the likelihood to choose the provider in the *Not Provided* framing than in the *Declined* framing (in which participants are already directed to consider the deliberate non-disclosure).

---

<sup>6</sup> Ten participants had response time durations 3 SD above the mean. Removing those participants do not change the pattern or significance of the results.

**Procedure.** Participants role-played patients searching for a doctor and were informed that all doctors in their region received a patient satisfaction rating ranging from 1 to 5 stars. They were presented with a description of what each star meant, with 1 indicating *very poor* and 5 indicating *excellent*, and then the profile of a doctor (eFigure 5). The final row of the profile was labeled, “Patient satisfaction rating” and participants read either that “Doctor Green did not provide the Rating” (*Not Provided* framing) or “Doctor Green declined to provide the Rating” (*Declined* framing).<sup>7</sup>

In the *Deliberation* condition, participants were first asked “In your opinion, why is the Patient Satisfaction Rating missing from Doctor Green’s profile?” and wrote their explanations. Participants were then asked to consider what the lowest Patient Satisfaction Rating Doctor Green would likely put on the profile. To do this, participants were encouraged to consider whether the doctor would be likely to disclosure a 5 star rating, or a 3 star or a 1 star rating. These two questions represented our perspective-taking manipulation. Thereafter, participants were presented with Doctor Green’s profile again and stated how likely they would be to choose the doctor on a 0 to 100 scale, and estimated the missing star rating.

Participants in the *Control* condition answered the same questions but in a different order. They first rated how likely they would be to choose Doctor Green and estimated the missing star rating. Only after this did they experience the perspective-taking manipulation: they wrote their explanations on why the Patient Satisfaction Rating was missing and they considered the lowest Patient Satisfaction Rating Doctor Green would likely put on the profile.

---

<sup>7</sup> The term “Declined” was used in place of “Refused” in Experiments 1 and 2, since we judged it a less loaded term that nonetheless conveyed clearly that it was the Doctor’s choice not to provide the information. “I decline to answer” is also used regularly particularly in legal contexts to refer to the right to not give information (Greer, 1990).

We coded participants' written explanations for why the Patient Satisfaction Rating was missing from Doctor Green's profile as explained below. First, we distinguished between those explanations that proposed non-deliberate versus deliberate withholding of information. Second, we distinguished between different types of "deliberate" explanations. Inter-coder reliabilities for each category, measured by Cohen's kappa, are shown in Table 1a:

- 1) *Non-deliberate*: This code was used when non-disclosure was explained as due to factors such as the rating not being available, technical difficulties, forgetting on the part of Doctor Green, or the high cost of displaying the rating. The essence of a non-deliberate explanation is that it suggests Doctor Green did not actively and intentionally withhold their rating. An example is, "*He seems like a capable doctor, perhaps he just forgot to put it on his profile.*" A non-deliberate explanation suggests the possibility of strategic naïveté.
- 2) *Deliberate*: This code was used when non-disclosure was interpreted as due to an active intentional choice by Dr. Green to withhold the information. The perspective-taking intervention in this study should move recipients to these "deliberate" categories by increasing deliberation on the non-disclosure (H2).
  - a. *Irrelevant* explanations were used when participants agreed the doctor had actively withheld their rating, but because the information was (a) of poor quality or unreliable, (b) irrelevant, or (c) ambiguous. While the category was broad, the core implication was that the reasons for non-disclosure were independent of the level of the rating. For example, "*Probably because these sorts of reviews tend to be awful for objective measure when people rate basic service a 5 or 1 bomb someone that looked at them wrong.*"
  - b. *Partial unraveling* implied that respondents suspected that the deliberate non-disclosure was due to the level of the rating and could have an effect on

consumer's evaluations. For example, respondents may have explicitly said that the ratings are poor, or that the doctor is "hiding something", or the ratings are bad for business. An example is "*He might either have a low score or just an average one.*" Although this code was used when recipients recognized non-disclosure as deliberate and reflecting something negative about the rating, it did not necessarily imply the ultimate level of unraveling implied by the game theory argument (that the doctor has the worst possible rating) thus we predicted recipients would still be charitable (H1b) despite the perspective-taking intervention.

- c. *Full unraveling* responses implied that participants gave the normatively sophisticated response as to why Doctor Green did not disclose his Patient Satisfaction Rating – in that, as predicted by game theory, it is because Dr. Green is likely to have the worst possible rating.
- d. *Other deliberate* responses indicated that the doctor had withheld the information, but either did not speculate as to why (for example, "*The doctor declined to provide the information*") or gave a reason that did not fit into the other categories (for example, "*maybe he already has enough patients and doesn't want to overload his current practice due to his high rating*").

Each response could be coded for multiple explanations and some participants gave multiple reasons such as, "*He either doesn't like the rating and feels it reflects poorly on him, or he's old and doesn't really know how to set up his profile.*" If responses did not fall into any of these categories, they were coded as "other." For example, some participants indicated they did not know or were unwilling to speculate about the reason for the information being missing.

## Results

**Choosing the provider.** As predicted, there was an interaction between Framing and Intervention,  $F(1, 705) = 9.48, p = .002, \eta_p^2 = .01$ , but no main effects of Framing,  $F(1, 705) = 1.46, p = .23$ , or Intervention,  $F(1, 705) = .14, p = .71$ . The interaction took the predicted form: In the low-salience *Not Provided* framing, participants were less likely to choose the doctor in the *Deliberation* ( $M = 35.72, SD = 21.48$ ) than in the *Control* condition ( $M = 41.73, SD = 23.50$ ),  $F(1, 705) = 6.01, p = .015, \eta_p^2 = .008$ . In the *Declined* framing, there was no effect of deliberation, ( $M = 38.98, SD = 24.39$  vs  $M = 34.25, SD = 23.06$ ),  $F(1, 705) = 3.64, p = .06$ . We also replicated our findings from Experiments 1 and 2: participants were more likely to choose the doctor in the *Not Provided* condition vs. the *Declined* condition (without deliberation):  $F(1, 705) = 9.51, p = .002, \eta_p^2 = .01$ , but this difference was eliminated with deliberation,  $F(1, 705) = 1.70, p = .19$ .

**Estimate of the missing information.** There was a main effect of Framing,  $F(1, 705) = 5.63, p = .02, \eta_p^2 = .008$ , and a Framing by Intervention interaction,  $F(1, 705) = 6.46, p = .011, \eta_p^2 = .009$ , but no main effect of Intervention,  $F(1, 705) = .002, p = .97$ . The main effect of Framing was estimates of the patient satisfaction rating were lower in the *Declined* framing ( $M = 2.49, SD = .87$ ) than in the *Not Provided* framing, ( $M = 2.65, SD = .81$ ). The estimates in the *Declined* framing were nonetheless much higher than the minimum of 1 star, indicating the charitability of recipients towards the missing information,  $t(353) = 32.21, p < .001$ .

The interaction was as expected. In the *Control* conditions, the more salient *Declined* framing ( $M = 2.42, SD = .87$ ) produced lower estimates than the *Not Provided* framing ( $M = 2.73, SD = .91$ ),  $F(1, 705) = 12.49, p < .001, \eta_p^2 = .02$ . But in the *Deliberation* conditions, there was no effect of framing ( $M = 2.58, SD = .86$  vs  $M = 2.57, SD = .72$ ),  $F(1, 705) = .01, p = .91$ .

**Lowest star rating the doctor would disclose.** The effect of perspective-taking, while significant, was smaller than would be expected if participants were fully strategic. An explanation for this is the strikingly high level of charitability in consumer's estimates of the doctor's personal breakeven point, the point at which the doctor will no longer disclose; the mean was 3.24 stars ( $SD = .68$ ), much higher than the lowest 1 star rating,  $t(708) = 69.43$ ,  $p < .001$ .<sup>8</sup> This means that even reflecting on reasons for non-disclosure does not produce maximum suspicion on the part of consumers.

**Coding and analyses of reasons for why information is missing.** Table 1a summarizes the number and percentage of reasons given that fall into the main categories of *Non-deliberate*, and *Deliberate* (*Irrelevant*, *Partial unraveling*, *Full unraveling* and *Other deliberate* explanations). Table 1b gives example statements for each category.

Our perspective-taking intervention seemed to work (i.e., increased deliberation on the non-disclosure) in that when asked to consider why the information was missing the great majority (86%) of participants provided at least one *Deliberate* explanation. However, there were responses in each category: 13% of recipients offered *Non-deliberate* explanations for non-disclosure, suggesting a degree of strategic naïveté. Of those who gave *Deliberate* explanations, 15% suggested the missing information might have been judged *Irrelevant*, and 75% inferred something negative about the rating, demonstrating at least *Partial unraveling*. Both these categories reveal some degree of charitability towards the doctor. Only one participant engaged in *Full unraveling* and contemplated that the doctor's rating is probably a 1 star (this participant was in the Declined framing condition).

---

<sup>8</sup> In all conditions, participants gave estimates of the lowest star rating the doctor would disclose after their written explanations of why the rating was missing (this was part of the perspective-taking deliberation manipulation); thus, we expected and found no differences between conditions (all  $ps > .15$ ).



The framing of non-disclosure as Declined vs. Not-provided had the predicted effect on whether explanations referred to *Non-deliberate* withholding of information. *Non-deliberate* explanations were more likely in the Not provided frame (19%) than in the Declined (8%) frame,  $X^2(1) = 20.30, p < .001$ . There was no effect of framing on the sub-categories of the *Deliberate* explanations: We suggest that while the Declined frame made the non-deliberate explanation less plausible, it did not make people more likely to view the missing information as negative.

## Discussion

Experiment 3 identified a perspective-taking intervention that also increased the salience of non-disclosure by emphasizing how it occurred deliberately. While the perspective-taking intervention did not lead to full unraveling and maximum skepticism (aside from one participant; H1b), participants did realize that the undisclosed information was likely to be deliberately withheld and less favorable for the doctor (H2). This demonstrates that interventions to increase perspective-taking can increase consumers' deliberation on the missing information and consequently become a bit more skeptical about non-disclosure.

### Experiment 4: Comparing Multiple Profiles and Highly Salient Non-Disclosure

One way deliberate non-disclosure can become salient is through comparison with disclosed information. If you compare the specifications for two cell phones, and one states the battery life and the other does not, you might infer that if the phone manufacturer did not mention battery life it is because the battery life is inadequate relative to other phones on the market. Studies that examine separate vs. joint (side-by-side) evaluations support the view that joint evaluations make comparisons easier by providing a reference to which one can compare and contrast (Hsee, 1996; Hsee & Zhang, 2010; Kardes & Sanbonmatsu, 1993). In

economics, much theoretical work has discussed how unraveling can occur when there is competition between providers, even when the consumer might not initially know the information exists (e.g., Milgrom, 2008).

In this experiment, we investigated responses to deliberate non-disclosure when it can be contrasted with disclosure by other parties. In addition to this “other party disclosure,” we again flagged deliberate non-disclosure, with different non-disclosure framings. We predicted that consumers would display less strategic naïveté and charitability if they were able to compare across providers (H2).

## Method

Students from a U.S. university ( $N = 117$ ; 47 women, 70 men;  $M_{age} = 21.7$ ,  $SD = 1.3$ )<sup>9</sup> evaluated four different doctor profiles from the perspective of patients searching for a doctor. Two key pieces of information were the doctors’ trustworthiness and conflicts of interest which were either disclosed or not disclosed. Trustworthiness ratings had a range from 30 (lowest possible) to 90 (top rating) and conflict of interest statements could be classified as none, minor or strong. All other information on the profile was disclosed. Patients evaluated all four physicians in any order – each physician profile was on a separate page, but participants could move freely back and forth between profiles to compare them for as long as they wanted. The profiles and conditions are described below (also see eFigure 6 in the supplement).

**Physician profiles.** Doctor W was our *baseline full disclosure* profile, i.e., the doctor always had both trustworthiness and conflict of interest statements disclosed. The specific levels of trustworthiness and conflict of interest were chosen randomly from the three possible levels for trustworthiness (50, 70, or 80), and for conflict of interest (none, minor or

---

<sup>9</sup> Three participants had response time durations 3 SD above the mean. Removing those participants do not change the pattern or significance of the results.

strong). Doctor X was our *doctor of interest* (ratings for Doctor X were our main dependent variable). This doctor had low (50) trustworthiness (but not the worst possible which was 30) and a strong conflict of interest. This information was *optional*, meaning in some conditions it was disclosed and in others not-disclosed.

The remaining two doctors, Doctors Y and Z, were comparison doctors, included to make the missing information from Doctor X's profile more conspicuous. Doctor Y had moderate trustworthiness (70) and no conflict of interest. Doctor Y's trustworthiness was *optional* information. Doctor Z had high trustworthiness (80) and a minor conflict of interest. Doctor Z's conflict of interest was *optional*.

**Conditions.** Participants were randomized into four between subject conditions, which differed in whether optional information was disclosed, and in how non-disclosure was flagged. In the *Full Disclosure* condition, all information was displayed for all four physician profiles. This condition provided us with a baseline for Doctor X, who had the worst possible value on conflict of interest and a trustworthiness level of 50.

In the remaining three conditions, the optional information was undisclosed, and varied as follows: The least salient *Absent* condition did not display the information, the table row or the name of the missing dimension (as in Experiment 2). The table row and information were present for other profiles, so participants would have to compare profiles to notice that information appeared in some but not other profiles. In the moderate salience *Blank* condition, the table row and name of the missing dimension, along with the range of possible values for the missing dimension were provided and there was a blank space where the missing information should be displayed. In the high salience *Declined* condition, as with the *Blank* condition, the table row, dimension name and range of values for the dimension were displayed and in the space where the missing information should be displayed,

participants saw “declined to disclose.” After comparing the profiles, participants indicated how likely they were to choose each of the four physicians on a scale from 0 to 100.

## Results and Discussion

**Choosing Doctor X.** We conducted a univariate ANOVA with the likelihood of choosing Dr. X as the dependent variable, and the four between subject conditions as the independent variable. Covariates included the likelihood of choosing Dr. W, Dr. W’s trustworthiness rating and Dr. W’s conflict of interest (rated on a 3-point scale, with 1=none, 2= minor, and 3=strong).<sup>10</sup> We conducted pairwise comparisons between the four disclosure conditions, specifically comparing *Full Disclosure* with the *Absent*, *Blank* and *Declined* non-disclosure conditions, and comparing the three non-disclosure conditions to each other.

Ratings for Dr. X displayed a significant main effect of disclosure condition,  $F(3, 110) = 16.06, p < .001, \eta_p^2 = .31$  (Figure 4). Pairwise comparisons revealed no significant difference ( $p = .19$ ) between the *Full Disclosure* condition ( $M = 44.82, SD = 23.58$ ) and the *Declined* condition ( $M = 48.04, SD = 28.21$ ) for the likelihood to choose Doctor X. Thus, we find that when multiple profiles were explicitly combined with the strong words “declined to disclose” participants did infer it to be equivalent to full disclosure of poor information.

The remaining non-disclosure conditions, *Blank* ( $M = 67.10, SD = 14.56$ ) and *Absent* ( $M = 69.52, SD = 16.04$ ), were not significantly different from each other ( $p = .48$ ) but produced significantly higher choice likelihoods than the *Full Disclosure* condition (both  $p$ ’s  $< .001$ ).

Overall, Experiments 1 through 4 all show some success in our interventions to induce people to treat undisclosed information a little more skeptically. Indeed, when participants viewed multiple profiles side by side when deliberate non-disclosure was made

---

<sup>10</sup> Results are similar if the covariates are removed from the model.

highly salient and flagged by the words “declined to provide,” participants displayed the maximum skepticism predicted by game theory.

### **Experiment 5: Providers’ Reasons for Deliberate Non-Disclosure**

In Experiment 5, we investigated what providers would do when given the opportunity to disclose or withhold information of varying quality about themselves, when consumers would know that they had made that decision. Participants simulated the role of physicians deciding what information to put on their website: they made disclosure decisions and justified those decisions. We predicted that just as the consumers in our previous studies showed limited strategic behavior, so, for the most part, would the providers in this study—i.e., they would not disclose at the rates predicted by unraveling theory. Importantly, the reasons they would give for their decisions would not typically contain strategic thinking (H3). We expected most providers would base their assessments on Level 0 thinking (Camerer & Johnson, 2004), in which they disclosed information they believed would give a good impression without considering the effects of not disclosing.

### **Methods**

**Participants, design and procedure.** Mturkers ( $N = 310$ ; 112 women, 197 men, 1 gender unreported;  $M_{age} = 31.2$ ,  $SD = 11.2$ )<sup>11</sup> simulated the role of a physician compiling an online profile to attract new patients and were randomized into one of nine conditions in a 3 (Trust: low vs. moderate vs. high) x 3 (Conflict of Interest: none vs. minor vs. strong) between subject design.

Each participant first viewed their main profile (see eFigure 7) which contained information that had to be on their profile, such as their specialty (internist), their admitting

---

<sup>11</sup> Three participants had response time durations 3 SD above the mean. Removing those participants do not change the pattern or significance of the results.

hospital, board certification, certification date, medical school, and faculty appointment. They were next presented with two optional profile items: residency information (a filler item) and a conflict of interest disclosure statement. Conflict of interest was randomized to three levels: strong (received over \$500,000 from Pfizer in 2010, 2011 and 2012), minor (received \$1,000 from Wyeth in 2010), and none (no competing interests). They chose whether to disclose both residency and conflict of interest.

Participants next viewed their average patient evaluation scores which included trustworthiness. They were told these evaluations came from patients who used scales from 0-100, where 0 is extremely poor and 100 is outstanding. The range of evaluations for all doctors in their region was also displayed. Three of the evaluations were optional to display: two filler items (office environment and friendliness—both at the top of the scale for all conditions) and the trustworthiness rating. Participants were randomized to have either low (50), intermediate (70), or high (80) trustworthiness ratings. The low rating was nowhere near the bottom of the possible range (30), and the high rating was not quite at the top (90).

Participants were informed their potential patients would “see the labels exactly as they are shown [on the profile], including the ranges, even if the information about you is left blank.” They therefore knew that patients would see an online profile similar to what participants viewed in the *Blank* condition of Experiments 2 and 4.

Participants also stated why they made their decision to reveal or not reveal their trustworthiness and conflict of interest. These responses were coded into the following four categories: (inter-coder reliabilities for each category, measured by Cohen’s kappa, are shown in Table 2a):

- 1) *Strategic* explanations include at least a minimal understanding that consumers would recognize non-disclosure and respond negatively to it. An example is: “If I had not included it, patients would be able to see a blank spot under Conflict of Interest

Disclosure. It is better to disclose it than have patients think I'm hiding something.”

When coding, we were very liberal in our interpretation of strategic thinking. To refer to the “Level-k” analysis discussed earlier, anything from “Level-1” or above would be classed as Strategic.

- 2) *Naïve inference* explanations implied that the information disclosed would have an effect on consumer’s evaluations, but with no reference to what would happen if it were not disclosed. An example was “If people see that my trustworthiness rating is 70 they may not trust me.” Such a statement, if not accompanied by (for instance) “... but if they don’t see the rating they may not trust me either” was coded as a naïve inference. With reference to the Level k analysis, this would amount to Level 0 thinking.
- 3) *Irrelevant* explanations were used to imply the information was (a) of poor quality or unreliable, (b) irrelevant, or (c) ambiguous. For example, “Because I don't think that [rating] would affect my judgement as a doctor, so I don't think it's relevant.” Providers often used these reasons in a way that we interpreted as ad-hoc justifications for not disclosing. Note that irrelevant explanations are not a “strategically rational” reason for not disclosing, because consumers should still notice that information is missing.
- 4) *Ethics*. These explanations referred to ethical obligations, usually to disclose information. Terms like “ought,” “should,” “honesty,” “fairness,” “integrity,” and “transparency” were associated with these explanations. An example response is: “I try not to hide much about myself in general, so I chose to display everything. I enjoy being as transparent as I can.”

If responses did not fall into any of these categories, they were coded as “other.”

We have not listed a category of response that might be called “sophisticated normative reasoning.” In line with how the term sophisticated is used in behavioral economics, this would involve explaining that non-disclosure is the best decision because consumers won’t think strategically. For instance, a participant might have said “If patients

were rational strategic thinkers I would disclose, but they are not.” We scoured all responses for hints of such reasoning but found none.

## Results

**Disclosure decisions.** As expected, disclosure decisions did not follow the pattern predicted by unraveling theory (see eFigure 8 in the supplement). According to theory, everyone would disclose their trustworthiness rating. But at a trustworthiness rating of 50 only 28% disclosed. Significantly more (54%) disclosed when the rating was 70,  $\chi^2(1) = 14.95, p < .001$ , and more again (88%) when it was 80,  $\chi^2(1) = 27.81, p < .001$ . In no group was there 100% disclosure. Similarly, in none of the conflict of interest groups was there 100% disclosure or even close to it. Significantly more participants disclosed no conflict (62%) than a minor conflict (42%),  $\chi^2(1) = 8.39, p = .004$ ; and more displayed minor conflicts than strong conflicts (20%),  $\chi^2(1) = 10.31, p = .001$ . The overall pattern is one of “partial unraveling,” in which the providers disclosed at a higher rate the better their performance but did not disclose anywhere near the rate predicted by unraveling theory.

**Reasons for (non)disclosure.** Table 2a summarizes the number and percentage of reasons given that fall into the four categories (*Strategic, Naïve inference, Irrelevant, and Ethical*) when making disclosure decisions under all levels of trustworthiness and conflict of interest. Table 2b gives example statements for each category.

Supporting H3, few participants provided even rudimentary *Strategic* explanations for their disclosure decisions (only 9% for trustworthiness disclosures and 4% for conflict of interest disclosures). The use of strategic reasoning was independent of the specific level of trustworthiness,  $\chi^2(2) = 1.95, p = .38$ , and conflict of interest,  $\chi^2(2) = 3.30, p = .19$ , and all but one who used strategic reasoning disclosed their information. The one exception was a participant who came to a strategic insight *after* their disclosure decision: “*I viewed 50 as bad and would bring negativity to my name. Looking back this may make me look less trustworthy*



*though.*” In addition, virtually everyone who used strategic reasoning for trustworthiness and conflict of interest disclosed their residency, office environment and friendliness information as well.

The *Irrelevant* category contained responses that expressed concerns that the information was irrelevant, unreliable or ambiguous, and usually was used as a justification for non-disclosure. This category was more widely used for the non-disclosure of conflict of interest statements (25%) than for trustworthiness (9%). Conflicts of interest were often described as being irrelevant or something that patients do not need to know, as in “... *it has nothing to do with my quality of care. It also is not really anyone’s business how I conduct myself as long as it is in best interests of the patients and legal.*” For trustworthiness, responses in this category typically revolved around issues of information reliability or accuracy, as in “*it wasn’t so low that it would put the average person off, but it’s kind of a weird metric. How do you measure that accurately?*”

The *Naïve inference* category was used far more frequently than any other category (66% for trustworthiness and 55% for conflict of interest). Providers chose based mainly on how they, and consumers, would directly evaluate the information if they saw it, but without considering what consumers might infer when information is missing (strategic naiveté). As Tables 2a and 2b show, this reasoning was mostly used to explain disclosure of high ratings, and non-disclosure of low ratings. The primary determinant of whether to disclose, therefore, was based on the (naïve) interpretation of the content being disclosed.

*Ethical* reasons were given 15% of the time for trustworthiness and 19% of the time for conflict of interest. Ethical reasoning was associated with a high level of disclosure, even when the information to disclose was not the best. Of the 31 participants who disclosed a low trustworthiness score, 14 (45%) gave an ethical explanation, as did 11 of the 19 participants

(58%) who disclosed a high conflict of interest. As seen in Table 2a, ethical explanations were the best predictor of disclosure for high conflicts of interest or low trustworthiness.

A few reasons fell into the “other” category (see Table 2b). One issue raised by participants with regards to conflict of interest is that by disclosing information you make people consider possibilities they otherwise would not. If people do not realize information is missing, it is a strategically valid reason for not disclosing it (Hirshleifer & Teoh, 2003). We had emphasized in the instructions that missing information would be highlighted on the doctor’s web site, but a (few) providers either ignored this instruction, or else felt consumers might not notice anyway.

## **Discussion**

As predicted (H3), there was limited strategic thinking by providers in this study. Only a few explanations reflected strategic reasoning, and those that did were almost invariably accompanied by disclosure. There was no evidence to suggest participants did not disclose because they thought consumers were naïve. Non-disclosure, at least in this setting, is not being used as a sophisticated strategic response (as previous theoretic work suggests: Brown et al., 2013; Hirshleifer & Teoh, 2003; King & Wallin, 1991). Instead, participants simply did not entertain any version of the unraveling argument.

The most common reason given for disclosing was to publicize good information and the most common reason for not disclosing was to withhold bad information. This is a manifestation of naïve “Level-0” thinking (e.g., Camerer & Johnson, 2004). The decision to not disclose, taken by most providers deciding whether to disclose average or poor information, demonstrate a lack of strategic thinking in that they failed to consider the responses of consumers to that missing information (H3). As our prior experiments show, however, consumers do not always draw negative inferences about missing information; thus,

the lack of strategic thinking by providers might be an (unforeseen) optimal response due to the corresponding naiveté and charitability of consumers.

### **Experiment 6: Medical Professionals' Reasons for Deliberate Non-Disclosure**

In this experiment, we examine the generalizability of the results from lay participants in Experiment 5 to investigate if actual physicians and health care professionals would behave similarly when deciding whether to disclose or withhold information of varying quality. These real providers viewed the same stimuli materials as in Experiment 5, made disclosure decisions and justified those decisions. We predicted that these real providers would behave similarly to our lay participants in Experiment 5 and base their disclosure decisions on what they believed would give a good impression to information recipients without considering the effects of not disclosing.

### **Methods**

**Participants, design and procedure.** Physicians and health care professionals ( $N = 56$ ; 34 women, 22 men;  $M_{age} = 41.6$ ,  $SD = 9.7$ )<sup>12</sup> from an executive Healthcare Leadership course at a U.S. business school took part in a similar study as that in Experiment 5. They imagined that they were compiling a physician profile and were randomized into one of four conditions in a 2 (Trust: low vs. high) x 2 (Conflict of Interest: none vs. strong) between subject design.

As in the prior experiment, each participant decided what optional information to disclose or not disclose on the physician profile. Unlike the prior experiment there were only two (rather than three) conflict of interest and trust levels (due to the smaller sample size of real providers that we had access to). The rest of the procedure including the coding was as described in the prior experiment.

---

<sup>12</sup> No participants had response time durations 3 SD above the mean.

## Results

**Disclosure decisions.** Disclosure decisions followed the same pattern as in the prior experiment: expert providers were less likely to disclose a low trustworthiness rating of 50 (15%) versus a higher rating of 80, (80%),  $\chi^2(1) = 23.26, p < .001$ , and they were less likely to disclose strong conflicts (29%) versus no conflict (66%),  $\chi^2(1) = 7.29, p = .007$ . Again, this pattern is one of partial unraveling, in which the providers disclosed at a higher frequency the better their ratings but they did not disclose anywhere near the rate predicted by unraveling theory.

**Reasons for (non)disclosure.** Table 2c summarizes the number and percentage of reasons given that fall into the four categories (*Strategic, Naïve inference, Irrelevant, and Ethical*) when making disclosure decisions under all levels of trustworthiness and conflict of interest.

As in the prior experiment, supporting H3, few providers gave an even rudimentary *Strategic* explanation for their disclosure decisions (only 7% for trustworthiness disclosures and 4% for conflict of interest disclosures). The use of strategic reasoning was again independent of the specific level of trustworthiness,  $\chi^2(1) = 0.80, p = .37$ , and conflict of interest,  $\chi^2(1) = 0.04, p = .85$ , and all but one who used strategic reasoning disclosed their information.

The *Irrelevant* category contained responses that expressed concerns that the information was irrelevant, unreliable or ambiguous, and again was mainly used as a justification for non-disclosure. This category was more widely used for the non-disclosure of conflict of interest statements (27%) than for trustworthiness (14%). Similar to our lay participants, real providers also described conflicts of interest as being irrelevant or something that patients do not need to know, as in “*Irrelevant to patient care - more important to disclose when conducting public education sessions or disclosing research for*

*collaborative interests/projects with pharmaceutical companies.”* For trustworthiness, providers mentioned issues of information accuracy and subjectivity, as in “*Trustworthiness is sort of a subjective virtue. Some patients may consider you not trustworthy simply because you charged higher than the next doctor. But they may not be aware that you are more experienced or maybe charge higher to keep a certain kind of clientele.*”

Like lay participants, real providers used the *Naïve inference* category far more frequently than any other category (73% for trustworthiness and 64% for conflict of interest). Providers chose based mainly on how they, and patients, would *directly* evaluate the information if they saw it, but without considering what patients might infer when information is missing. As Table 2c shows, this reasoning was mostly used to explain disclosure of high ratings, and non-disclosure of low ratings (note that some of the providers thought a rating of 80 for trustworthiness was not “high enough” for a physician practice – recall that the top rating for physicians in their region is 90 -- e.g., “*A higher score is expected in the public opinion to promote the office*”). Once again, therefore, the crucial determinant of whether to disclose was based on the direct inference of the content being disclosed.

*Ethical* reasons were given 11% of the time for trustworthiness and 12% of the time for conflict of interest. Ethical reasoning was almost always associated with disclosure regardless of whether the information was good or poor. Of the four providers who disclosed a low trustworthiness score, three (75%) gave an ethical explanation, as did six of the seven providers (86%) who disclosed a strong conflict of interest. Thus, ethical explanations were the best predictor of disclosure for high conflicts of interest or low trustworthiness.

A few reasons fell into the “other” category. Like lay participants, real providers also raised the issue that even mentioning the term conflict of interest may make patients worry

about them even if you had none. Real providers also noted that conflict of interest information is already available to patients through the Open Payments / Sunshine Act (see <https://www.cms.gov/openpayments/>) and thus not necessary to disclose. Although we had emphasized in the instructions that missing information would be highlighted on the doctor's web site, a few providers either ignored this instruction, or else felt patients might not notice anyway.

## **Discussion**

This experiment reveals that real expert providers behaved similarly to lay participants in Experiment 5. Again, as predicted (H3), only a very few explanations reflected strategic reasoning. The most common reason given for disclosing was to publicize good information and the most common reason for not disclosing was to withhold bad information consistent with "Level-0" thinking.

## **General Discussion**

We find evidence of limited strategic thinking about deliberate non-disclosure by both information providers and recipients. Recipients demonstrate strategic naiveté when they do not distinguish between deliberate and non-deliberate non-disclosures (Experiment 1). Even when they notice that disclosure is deliberate, they show high levels of charitability towards the missing information and to the information provider (Experiments 2 to 4).

Interventions to increase the strategic thinking of recipients had varying success. Indeed, recipients were almost completely unresponsive to non-disclosure unless their attention was drawn to it by terms such as "declined" or "refused." Recipients were more likely to note non-disclosure was deliberate, and they exhibited less charitability, when they considered why a provider might not disclose (Experiment 3). They were also more likely to

notice deliberate non-disclosure when they could compare multiple providers, some of whom disclosed and some of whom did not (Experiment 4).

Like recipients, providers demonstrated strategic naiveté. Providers mainly considered how consumers would interpret *disclosed* information and they did not even entertain the possibility that consumers could make inferences about non-disclosed information (Experiments 5 and 6).

### **An Analysis of the Assumptions for Full Unraveling**

Game theory predicts that strategically rational consumers will infer that deliberate non-disclosure of decision-relevant information means that information is the worst possible, providing certain assumptions are met. These assumptions (outlined more extensively in the introduction) include unambiguous ranking of dimensions, and costless and credible disclosures (i.e., the provider has the information and cannot lie). In the “outside world” it is unlikely these assumptions hold all the time. Considering this, non-disclosure may be the optimal choice for providers even if consumers are strategically rational. And if some (even high-quality) providers do not disclose when they “should,” then it becomes increasingly possible for other providers (especially mediocre and poor-quality ones) who might otherwise disclose to hide themselves in the thicket of non-disclosers. Consequently, full unraveling may not occur even if both providers and consumers are strategically rational.

One domain in which the assumptions for full unraveling may not be met is in the disclosure of calorie counts in restaurant menus. First, the ranking is ambiguous since fewer calories in a meal are not necessarily better than more calories (Downs, Wisdom, Wansink, & Loewenstein, 2013). It may also be costly to acquire calorie count information for each dish displayed on the menu. Moreover, knowledge of calories can also decrease consumers’ enjoyment of food (Thunström, 2019; Thunström, Nordström, Shogren, Ehmke, & van’t Veld, 2016; Woolley & Risen, 2018), and this imposes a further cost of disclosure in the form

of consumer dissatisfaction. We should therefore not expect universal disclosure of calorie counts. Indeed, this is why legislation is being used to require restaurants to provide them (Bollinger, Leslie, & Sorensen, 2011). And that legislation is itself controversial in part because it may not benefit consumers (e.g., Dumanovsky et al., 2011; Elbel, Kersh, Brescoll, & Dixon, 2009; Kiszko, Martinez, Abrams, & Elbel, 2014; Swartz, Braxton, & Viera, 2011).

Another situation in which the unraveling assumptions will not fully apply is when an information provider has a vast amount of information and it is unfeasible to disclose all of it. Too many disclosures can overwhelm consumers, as can lengthy, complex or unintelligible disclosures (Ben-Shahar & Schneider, 2011). A rational provider will want to make selective disclosure and so consumer skepticism should be tempered by this consideration. Other psychological factors can also come into play. For instance, too much positive information can lead consumers to suspect selective disclosure so that withholding even positive information can sometimes make sense (Campbell & Kirmani, 2000; Shu & Carlson, 2014).

Sometimes the assumptions for full unraveling appear not to be met but actually are. Consider, for instance, the case of grade non-disclosure policies seen at some schools, particularly U.S. business schools. This policy can prohibit universities from revealing student grades or graduate GPAs to recruiters. Despite this prohibition, students are free to inform recruiters about their undergraduate GPAs or their business school test scores, honors, or awards, such as the Dean's List. Likewise, employers are often forbidden from asking potential employees for personal information, such as health status. However, job candidates can voluntarily provide this, or similar "proxy", information, and so employers can draw inferences about the failure to make a voluntary provision (Peppet, 2011). As Richard Posner (1998, p. 103) put it: "it is difficult to see how a pooling equilibrium is avoided in which privacy is 'voluntarily' surrendered, making the legal protection of privacy futile." We need



not be quite as despairing as Posner to see that the *opportunity* to voluntarily provide information can sometimes turn into an *obligation* to provide that information.

As we have shown, however, the strict theoretical approach to unraveling is only the starting point to understanding what will happen when some people have the opportunity to disclose, and others have the opportunity to interpret that (non-) disclosure. Even when unraveling assumptions are met, both our experiments and evidence from the field (Brown et al., 2012; Jin & Leslie, 2003; Mathios, 2000) demonstrates that consumers do not infer the worse from missing information. Consequently, providers can “get away” with non-disclosure. Indeed, if we allow for strategically naïve or charitable consumers, there will be incomplete disclosure by providers.

In many cases, non-disclosure may be a preferred strategy for providers because the worst inferences are still in the imagination of consumers rather than in stark reality (Frank, 2008). That is, there is a difference between what can be inferred from the absence of information, and the concrete fact being revealed. A (guilty) suspect who refuses to testify might reasonably assume that her guilt-affirming silence will have less impact on a jury than an outright confession (Kassin, 2005; Kassin & Sukel, 1997). Indeed, as we found in Experiments 1 to 4, simply informing people that information that could have been disclosed but was not barely influenced consumers. But as the non-disclosure became more salient and closer to an uncooperative scenario, as indicated by the provider having “refused” or “declined” to provide information, people became increasingly wary of the provider. This inference could explain why John et al. (2016) found full unraveling in their studies when recipients were informed that a provider had “chosen not to answer” a question. A refusal to answer a question increases the salience of that deliberately withheld information.

### **Is Less Disclosure Sometimes Better?**

It is important to consider the consequences to consumers of full disclosure, however it is achieved. The effects are not always positive. For example, if people do not want information that is negative, they may choose to avoid seeing it even if it might improve their decision-making (Gigerenzer & Garcia-Retamero, 2017; Golman, Hagmann, & Loewenstein, 2017; Sweeny, Melnyk, Miller, & Shepperd, 2010). Research on conflict of interest disclosures show multiple ways in which such disclosures can produce unintended consequences for consumers (Loewenstein, Sah, & Cain, 2012; Sah, 2016). For example, conflict of interest disclosures can increase pressure to comply with poor-quality advice due to consumers' reluctance to signal their distrust in their provider (Sah, Loewenstein, & Cain, 2013, 2019). Conflict of interest disclosures might also paradoxically increase trust in the provider's expertise, especially when consumers are distracted and process the disclosure automatically rather than deliberatively (Sah, Malaviya, & Thompson, 2018). The opposite effect may also occur, in which disclosures make consumers skeptical and therefore prone to reject even high quality advice (Sah & Feiler, 2019). Conflict of interest disclosure can also affect providers themselves, increasing or decreasing bias in their advice (Sah, 2019).

In summary, full disclosure may sometimes have unintended consequences. However, for information that is important, reliable and relevant to a decision (and in conditions in which the assumptions for full unraveling are met), our studies show that consumers are insufficiently strategic. Consumers do not recognize when information has been withheld and this gives information providers an opportunity to withhold information they do not want consumers to have, knowing that those consumers are unlikely to punish them.

### **Practical and Policy Implications**

In addition to our theoretical contributions, our research has practical implications that organizations and policy makers might consider. The analysis of whether disclosure improves

consumer welfare is complicated and the difficulty is aggravated when consumers have shifting preferences (Sunstein, 2018). If the goal is to increase human welfare, careful decisions must be made about what to disclose and what not to disclose.

Some disclosures make consumers feel unhappy or unsatisfied, such as calories on restaurant menus (Thunström, 2019) or knowledge of the date of your partner's death (Gigerenzer & Garcia-Retamero, 2017). But there is also the instrumental value of knowing certain information even if it makes you unhappy (Adler, 2012). For example, you might make better choices after knowing the result of a medical test.

For the experiments in this paper (consumers researching profiles of doctors) and in other domains (buying a car or computer, going to a restaurant or deciding on a movie), it is likely that consumers are searching for information that is important and relevant to them to make a decision. In these cases, when relevant information disclosure is an important policy goal, voluntary disclosure is unlikely to be an adequate substitute for mandatory disclosure, even if the conditions for unraveling are met. If mandatory disclosure cannot be imposed, then some way of communicating that information has been deliberately withheld should be sought. Disclosure, therefore, may be effective not merely because of what is disclosed, but also by alerting consumers to the decisions of those who have chosen not to disclose. The latter goal can be achieved by giving providers the *opportunity* to disclose and informing consumers that this opportunity was provided and declined. The findings from our studies suggest that helping consumers become savvy by strongly framing non-disclosures may go some way toward achieving the same goal as mandatory disclosure, particularly if consumers can compare disclosures and non-disclosures in multiple providers. Finally, as with all policy proposals, we recommend testing in the field to see if the policy has the intended beneficial effect.

**Conclusions**

Providers often have a choice about whether to disclose privately held information about the quality of their product or service. They can withhold information if the product or service quality is not good, but if they do, consumers may make inferences about the absence of information. Our findings, however, suggest that consumers are highly charitable regarding the absence of deliberately withheld information and only react negatively when non-disclosure is made highly salient, by either embedding non-disclosures among disclosures, or by emphasizing that the provider made a deliberate choice to conceal information. Given consumers' strategically naïve and charitable attitude toward non-disclosure, the most robust solution to protect and/or inform consumers is mandatory disclosure of crucial information.

### References

- Adler, M. (2012). *Well-being and fair distribution: Beyond cost-benefit analysis*. Oxford University Press.
- Benndorf, V. (2018). Voluntary disclosure of private information and unraveling in the market for lemons: An experiment. *Games*, 9(2), 23–40.
- Benndorf, V., Kübler, D., & Normann, H.-T. (2017). Depth of reasoning and information revelation: An experiment on the distribution of k-levels. *International Game Theory Review*, 19(04), 1750021.
- Ben-Shahar, O., & Schneider, C. E. (2011). The failure of mandated disclosure. *University of Pennsylvania Law Review*, 159(3), 647–749.
- Bollinger, B., Leslie, P., & Sorensen, A. (2011). Calorie posting in chain restaurants. *American Economic Journal: Economic Policy*, 3(1), 91–128.
- Brown, A. L., Camerer, C. F., & Lovallo, D. (2012). To review or not to review? Limited strategic thinking at the movie box office. *American Economic Journal: Microeconomics*, 4(2), 1–26.
- Brown, A. L., Camerer, C. F., & Lovallo, D. (2013). Estimating Structural Models of Equilibrium and Cognitive Hierarchy Thinking in the Field: The case of withheld movie critic reviews. *Management Science*, 59(3), 733–747.
- Butler, D., & Read, D. (2017). Unraveling and strategic disclosure: Evidence from the hospitality industry. *Working Paper*. <https://doi.org/dx.doi.org/10.2139/ssrn.3068601>
- Camerer, C. F., & Johnson, E. J. (2004). Thinking about attention in games: Backward and forward induction. In *In Brocas and Carillo (Eds.) The Psychology of Economic Decisions* (Vol. 2, pp. 111–129). Oxford University Press.

- Campbell, M. C., & Kirmani, A. (2000). Consumers' use of persuasion knowledge: The effects of accessibility and cognitive capacity on perceptions of an influence agent. *Journal of Consumer Research*, 27(1), 69–83.
- Carrillo, P., Cellini, S. R., & Green, R. K. (2013). School quality and information disclosure: Evidence from the housing market. *Economic Inquiry*, 51(3), 1809–1828.
- Downs, J. S., Wisdom, J., Wansink, B., & Loewenstein, G. (2013). Supplementing menu labeling with calorie recommendations to test for facilitation effects. *American Journal of Public Health*, 103(9), 1604–1609.
- Dranove, D., & Jin, G. Z. (2010). Quality disclosure and certification: Theory and practice. *The Journal of Economic Literature*, 48(4), 935–963.
- Dumanovsky, T., Huang, C. Y., Nonas, C. A., Matte, T. D., Bassett, M. T., & Silver, L. D. (2011). Changes in energy content of lunchtime purchases from fast food restaurants after introduction of calorie labelling: Cross sectional customer surveys. *British Medical Journal*, 343, d4464.
- Elbel, B., Kersh, R., Brescoll, V. L., & Dixon, L. B. (2009). Calorie labeling and food choices: A first look at the effects on low-income people in New York City. *Health Affairs*, 28(Suppl1), w1110–w1121.
- Forsythe, R., Isaac, R. M., & Palfrey, T. R. (1989). Theories and tests of "blind bidding" in sealed-bid auctions. *The Rand Journal of Economics*, 20(2), 214–238.
- Frank, R. H. (2008, April 13). To Disclose or Not? Ask the Frogs. *The New York Times*. Retrieved from <https://www.nytimes.com/2008/04/13/business/13view.html>
- Fung, A., Graham, M., & Weil, D. (2007). *Full disclosure: The Perils and Promise of Transparency*. New York: Cambridge University Press.

- Garcia-Retamero, R., & Rieskamp, J. (2009). Do people treat missing information adaptively when making inferences? *The Quarterly Journal of Experimental Psychology*, 62(10), 1991–2013.
- Gigerenzer, G., & Garcia-Retamero, R. (2017). Cassandra's regret: The psychology of not wanting to know. *Psychological Review*, 124(2), 179.
- Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research*, 42(3), 237–288.
- Golman, R., Hagmann, D., & Loewenstein, G. (2017). Information avoidance. *Journal of Economic Literature*, 55(1), 96–135.
- Goodman, J. K., Cryder, C. E., & Cheema, A. (2012). Data Collection in a Flat World: Strengths and weaknesses of mechanical turk samples. *Journal of Behavioral Decision Making*, 26(3), 213–224. <https://doi.org/10.1002/bdm.1753>
- Grande, D., Shea, J. A., & Armstrong, K. (2012). Pharmaceutical industry gifts to physicians: Patient beliefs and trust in physicians and the health care system. *Journal of General Internal Medicine*, 27(3), 274–279.
- Gray, R. (2013, June 28). Row over surgeons' mortality rates as those who withheld data are named. *The Telegraph*. Retrieved from <http://www.telegraph.co.uk/news/health/news/10149303/Row-over-surgeons-mortality-rates-as-those-who-withheld-data-are-named.html>
- Greer, S. (1990). The Right to Silence: A Review of the Current Debate. *Mod. L. Rev.*, 53, 709.
- Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics*, 24(3), 461–483.

- Grossman, S. J., & Hart, O. D. (1983). An analysis of the principal-agent problem. *Econometrica*, 51(1), 7–45.
- Hanauer, D. A., Zheng, K., Singer, D. C., Gebremariam, A., & Davis, M. M. (2014). Public awareness, perception, and use of online physician rating sites. *The Journal of the American Medical Association*, 311(7), 734–735.
- Hernandez, J. M. C., Han, X., & Kardes, F. R. (2014). Effects of the perceived diagnosticity of presented attribute and brand name information on sensitivity to missing information. *Journal of Business Research*, 67(5), 874–881.
- Hirshleifer, D., & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics*, 36(1), 337–386.
- Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes*, 67(3), 247–257.
- Hsee, C. K., & Zhang, J. (2010). General evaluability theory. *Perspectives on Psychological Science*, 5(4), 343–355.
- Jin, G. Z. (2005). Competition and disclosure incentives: An empirical study of HMOs. *RAND Journal of Economics*, 36(1), 93–112.
- Jin, G. Z., & Leslie, P. (2003). The effect of information on product quality: Evidence from restaurant hygiene grade cards. *The Quarterly Journal of Economics*, 118(2), 409–451.
- Jin, G. Z., Luca, M., & Martin, D. (2015). *Is no news (perceived as) bad news? An experimental investigation of information disclosure*. National Bureau of Economic Research.
- John, L. K., Barasz, K., & Norton, M. I. (2016). Hiding personal information reveals the worst. *Proceedings of the National Academy of Sciences*, 113(4).



- Johnson, R. D., & Levin, I. P. (1985). More than meets the eye: The effect of missing information on purchase evaluations. *Journal of Consumer Research*, 12(2), 169–177.
- Jones, I. R., Gabe, J., & Exworthy, M. (2013, June 28). Will publishing surgeon performance improve outcomes? *Health Service Journal*. Retrieved from <http://www.hsj.co.uk/comment/will-publishing-surgeon-performance-improve-outcomes/5060344.fullarticle>
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Kardes, F. R., Posavac, S. S., & Cronley, M. L. (2004). Consumer inference: A review of processes, bases, and judgment contexts. *Journal of Consumer Psychology*, 14(3), 230–256.
- Kardes, F. R., Posavac, S. S., Silvera, D., Cronley, M. L., Sanbonmatsu, D. M., Schertzer, S., ... Chandrashekar, M. (2006). Debiasing omission neglect. *Journal of Business Research*, 59(6), 786–792.
- Kardes, F. R., & Sanbonmatsu, D. M. (1993). Direction of comparison, expected feature correlation, and the set-size effect in preference judgment. *Journal of Consumer Psychology*, 2(1), 39–54.
- Kassin, S. M. (2005). On the psychology of confessions: Does innocence put innocents at risk? *American Psychologist*, 60(3), 215.
- Kassin, S. M., & Sukel, H. (1997). Coerced confessions and the jury: An experimental test of the "harmless error" rule. *Law and Human Behavior*, 21(1), 27.
- King, R. R., & Wallin, D. E. (1991). Market-induced information disclosures: An experimental markets investigation. *Contemporary Accounting Research*, 8(1), 170–197.

- Kiszko, K. M., Martinez, O. D., Abrams, C., & Elbel, B. (2014). The influence of calorie labeling on food orders and consumption: A review of the literature. *Journal of Community Health, 39*(6), 1248–1269.
- Loewenstein, G., Sah, S., & Cain, D. M. (2012). The unintended consequences of conflict of interest disclosure. *The Journal of the American Medical Association, 307*(7), 669–670.
- Luca, M., & Smith, J. (2015). Strategic disclosure: The case of business school rankings. *Journal of Economic Behavior & Organization, 112*, 17–25.
- Mariotti, M. (1995). Is Bayesian rationality compatible with strategic rationality? *The Economic Journal, 105*(432), 1099–1109.
- Mathios, A. D. (2000). Impact of Mandatory Disclosure Laws on Product Choices: An Analysis of the Salad Dressing Market. *The Journal of Law and Economics, 43*, 651.
- Milgrom, P. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics, 380–391*.
- Milgrom, P. (2008). What the Seller Won't Tell You: Persuasion and disclosure in markets. *Journal of Economic Perspectives, 22*(2), 115–131.
- Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review, 85*(5), 1313–1326.
- Peppet, S. R. (2011). Unraveling privacy: The personal prospectus and the threat of a full-disclosure future. *Northwestern University Law Review, 105*, 1153.
- Posner, R. A. (1998). Privacy. In P. Newman, ed. *The New Palgrave Dictionary of Economics and the Law*. U.K.: Grove Dictionaries.
- Rand, D. G. (2012). The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. *Journal of Theoretical Biology, 299*, 172–179.

Rose, S., Sah, S., Raed, D., Schmidt, C., Mercer, M., Michum, A., ... Robertson, C. (2019).

Patient responses to physician disclosures of industry conflicts of interest: A randomized field experiment. *Organizational Behavior and Human Decision Processes*.

Ross Jr, W. T., & Creyer, E. H. (1992). Making inferences about missing information: The effects of existing information. *Journal of Consumer Research*, 19(1), 14–25.

Sah, S. (2012). Conflicts of Interest and Your Physician: Psychological processes that cause unexpected changes in behavior. *The Journal of Law, Medicine & Ethics*, 40(3), 482–487.

Sah, S. (2016, July 8). The Paradox of Disclosure. *The New York Times*. Retrieved from [http://www.nytimes.com/2016/07/10/opinion/sunday/the-paradox-of-disclosure.html?\\_r=0](http://www.nytimes.com/2016/07/10/opinion/sunday/the-paradox-of-disclosure.html?_r=0)

Sah, S. (2019). Conflict of interest disclosure as a reminder of professional norms: Clients first! *Organizational Behavior and Human Decision Processes*, 154, 62-79.

Sah, S., & Feiler, D. (2019). Conflict of interest disclosure with high quality advice: The disclosure penalty and the altruistic signal. *Psychology, Public Policy and Law*.

Sah, S., & Loewenstein, G. (2014). Nothing to declare: Mandatory and voluntary disclosure leads advisors to avoid conflicts of interest. *Psychological Science*, 25(2), 575–584.

Sah, S., Loewenstein, G., & Cain, D. M. (2013). The burden of disclosure: Increased compliance with distrusted advice. *Journal of Personality and Social Psychology*, 104(2), 289–304.

Sah, S., Loewenstein, G., & Cain, D. M. (2019). Insinuation anxiety: Concern that advice rejection will signal distrust after conflict of interest disclosures. *Personality and Social Psychology Bulletin*, 45(7), 1099–1112.

<https://doi.org/10.1177/0146167218805991>

- Sah, S., Malaviya, P., & Thompson, D. (2018). Conflict of interest disclosure as an expertise cue: Differential effects due to automatic versus deliberative processing. *Organizational Behavior and Human Decision Processes*, 147, 127–146.
- Sanbonmatsu, D. M., Kardes, F. R., & Herr, P. M. (1992). The role of prior knowledge and missing information in multiattribute evaluation. *Organizational Behavior and Human Decision Processes*, 51(1), 76–91.
- Sanbonmatsu, D. M., Kardes, F. R., Houghton, D. C., Ho, E. A., & Posavac, S. S. (2003). Overestimating the importance of the given information in multiattribute consumer judgment. *Journal of Consumer Psychology*, 13(3), 289–300.
- Sanbonmatsu, D. M., Kardes, F. R., Posavac, S. S., & Houghton, D. C. (1997). Contextual influences on judgment based on limited information. *Organizational Behavior and Human Decision Processes*, 69(3), 251–264.
- Sanbonmatsu, D. M., Kardes, F. R., & Sansone, C. (1991). Remembering less and inferring more: Effects of time of judgment on inferences about unknown attributes. *Journal of Personality and Social Psychology*, 61(4), 546–554.
- Selten, R. (1988). *Models of strategic rationality*. Springer Science & Business Media.
- Shu, S. B., & Carlson, K. A. (2014). When three charms but four alarms: Identifying the optimal number of claims in persuasion settings. *Journal of Marketing*, 78(1), 127–139.
- Simmons, C. J., & Lynch, J. G. (1991). Inference effects without inference making? Effects of missing information on discounting and use of presented information. *Journal of Consumer Research*, 17(4), 477–491.
- Stahl II, D. O., & Wilson, P. W. (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior & Organization*, 25(3), 309–327.
- Sunstein, C. R. (2018). *The cost-benefit revolution*. MIT Press.

- Swartz, J. J., Braxton, D., & Viera, A. J. (2011). Calorie menu labeling on quick-service restaurant menus: An updated systematic review of the literature. *International Journal of Behavioral Nutrition and Physical Activity*, 8(1), 135.
- Sweeny, K., Melnyk, D., Miller, W., & Shepperd, J. A. (2010). Information avoidance: Who, what, when, and why. *Review of General Psychology*, 14(4), 340–353.
- Thom, D. H., Kravitz, R. L., Bell, R. A., Krupat, E., & Azari, R. (2002). Patient trust in the physician: Relationship to patient requests. *Family Practice*, 19(5), 476–483.
- Thunström, L. (2019). Welfare effects of nudges: The emotional tax of calorie menu labeling. *Judgment and Decision Making*, 14(1), 11.
- Thunström, L., Nordström, J., Shogren, J. F., Ehmke, M., & van't Veld, K. (2016). Strategic self-ignorance. *Journal of Risk and Uncertainty*, 52(2), 117–136.
- Viscusi, W. K. (1978). A note on "lemons" markets with quality certification. *The Bell Journal of Economics*, 277–279.
- Woolley, K., & Risen, J. L. (2018). Closing your eyes to follow your heart: Avoiding information to protect a strong intuitive preference. *Journal of Personality and Social Psychology*, 114(2), 230–245.

**Table 1a: Coding of reasons for why information is missing in Experiment 3**

Framing	N	Non-deliberate N (%)	Deliberate				Other N (%)
			Irrelevant N (%)	Partial unraveling N (%)	Full unraveling N (%)	Other deliberate N (%)	
Not Provided	355	68 (19)	50 (14)	261 (74)	0 (0)	10 (3)	11 (3)
Declined	354	27 (8)	54 (15)	273 (77)	1 (0)	6 (2)	6 (2)
<b>Total</b>	<b>709</b>	<b>95 (13)</b>	<b>104 (15)</b>	<b>534 (75)</b>	<b>1 (0)</b>	<b>16 (2)</b>	<b>17 (2)</b>
<b>Kappa</b>		<b>.98</b>	<b>.99</b>	<b>.98</b>	<b>1.00</b>	<b>.98</b>	<b>.99</b>

Note: One response can fall into multiple categories, and some participants received no code in any of these categories.

**Table 1b: Examples for the categories of missing information reasons in Experiment 3**

<b>Non-deliberate</b> (not available, technical, forgot, costly)	He seems like a capable doctor, perhaps he just <i>forgot</i> to put it on his profile There is <i>not enough data</i> to provide a rating. It could be just that <i>he missed the filing date or something like that</i> . The website might be new or maybe he's a younger doctor.
<b>Deliberate</b>	
<b>Irrelevant</b> (unreliable, irrelevant, inaccurate, ambiguous or 'right' not to disclose)	I think Doctor Green didn't really see the need to include the ratings. Maybe he doesn't think the ratings are accurately measured He did not feel that the ratings were a relevant reflection of his practice It is possible that he does not feel that they are accurately representative and wants people to decide for themselves
<b>Partial Unraveling</b> (the information is bad)	He might either have a low score or just an average one Possibly too low of a rating to talk about and rather just keep it private I find it very odd. It makes me think Doctor Green is hiding something It could be that he is unsatisfied with his rating and does not want people to judge him based on it
<b>Full Unraveling</b> (the information is the worst possible)	He probably has a rating of one
<b>Other deliberate</b>	The doctor declined to provide the information It could be deliberate

**Table 2a: Coding of reasons to disclose trustworthiness and conflict of interest statements in Experiment 5**

Disclosed	Trustworthiness	N	Strategic N (%)	Naïve Inference N (%)	Irrelevant N (%)	Ethical N (%)	Other N (%)
No	Low	81	1 (1)	64 (79)	15 (19)	0 (0)	3 (4)
	Intermediate	41	0 (0)	35 (85)	7 (17)	2 (5)	2 (5)
	High	13	0 (0)	5 (39)	4 (31)	2 (15)	3 (23)
	<b>Total</b>	<b>135</b>	<b>1 (1)</b>	<b>104 (77)</b>	<b>26 (19)</b>	<b>4 (3)</b>	<b>8 (6)</b>
Yes	Low	31	10 (32)	4 (13)	1 (3)	14 (45)	4 (13)
	Intermediate	49	11 (22)	23 (47)	0 (0)	15 (31)	3 (6)
	High	95	7 (7)	73 (77)	0 (0)	12 (13)	8 (8)
	<b>Total</b>	<b>175</b>	<b>28 (16)</b>	<b>100 (57)</b>	<b>1 (1)</b>	<b>41 (23)</b>	<b>15 (9)</b>
	<b>Total</b>	<b>310</b>	<b>29 (9)</b>	<b>204 (66)</b>	<b>27 (9)</b>	<b>45 (15)</b>	<b>23 (7)</b>
	<b>Kappa</b>		<b>.98</b>	<b>.81</b>	<b>.84</b>	<b>.87</b>	
Disclosed	Conflict of Interest	N	Strategic N (%)	Naïve Inference N (%)	Irrelevant N (%)	Ethical N (%)	Other N (%)
No	Strong	75	0 (0)	62 (83)	12 (16)	3 (4)	4 (5)
	Minor	54	0 (0)	34 (63)	31 (57)	2 (4)	1 (2)
	None	47	0 (0)	11 (23)	34 (72)	1 (2)	4 (9)
	<b>Total</b>	<b>176</b>	<b>0 (0)</b>	<b>107 (61)</b>	<b>77 (44)</b>	<b>6 (3)</b>	<b>9 (5)</b>
Yes	Strong	19	1 (5)	4 (21)	0 (0)	11 (58)	4 (21)
	Minor	39	5 (13)	10 (26)	0 (0)	22 (56)	3 (8)
	None	76	7 (9)	49 (65)	0 (0)	20 (26)	6 (8)
	<b>Total</b>	<b>134</b>	<b>13 (10)</b>	<b>63 (47)</b>	<b>0 (0)</b>	<b>53 (40)</b>	<b>13 (10)</b>
	<b>Total</b>	<b>310</b>	<b>13 (4)</b>	<b>170 (55)</b>	<b>77 (25)</b>	<b>59 (19)</b>	<b>22 (7)</b>
	<b>Kappa</b>		<b>.99</b>	<b>.81</b>	<b>.84</b>	<b>.87</b>	

Note: One response can fall into multiple categories.



**Table 2b: Examples for the categories of disclosure reasons in Experiment 5**

<b>Strategic</b> (rational disclosure / unraveling theory)	If I had not included it, patients would be able to see a blank spot under Conflict of Interest Disclosure. It is better to disclose it than have patients think I'm hiding something. <i>(Conflict of interest)</i>
	If the heading is going to be there than might as well let them see it so they don't think you are trying to hide something. <i>(Conflict of interest)</i>
	If I show the other two, but exclude that one, it will look untrustworthy and reaffirm my poor score. If it is hidden, people will probably assume it is even worse than it is. <i>(Trust)</i>
<b>Naïve Inference</b> (if the information is good / bad)	Because some patients might be turned off and think of me as a "drug pusher." <i>(Conflict of interest)</i>
	I made that decision because I was afraid of the way that it would reflect upon me. <i>(Conflict of interest)</i>
	I would want people to know that a high percentage of patients thought I was trustworthy, to ease the mind of future patients. <i>(Trust)</i>
<b>Irrelevant</b> (unreliable, irrelevant, inaccurate or ambiguous)	The score sucks, a little more than half of the high end. I'm worried I would lose patients. <i>(Trust)</i>
	The information doesn't seem relevant. <i>(Conflict of interest)</i>
	I didn't think it would be relevant or helpful to someone look for potential doctors. <i>(Conflict of interest)</i>
	The rating could be skewed due to sample size. <i>(Trust)</i>
<b>Ethical</b> (honest, fair, transparent)	I feel my trust is higher than a 70 rating. <i>(Trust)</i>
	I believe that honesty is the best policy. <i>(Conflict of interest)</i>
	I think it's prudent and ethical for a doctor to list his conflicts of interests. <i>(Conflict of interest)</i>
	I want to be as transparent as possible. <i>(Trust)</i>
<b>Other</b>	I want to be completely transparent. People should have all the information they need to make an informed decision. <i>(Trust)</i>
	I think having it on my profile would make people think about it even though I don't have any conflict of interest. <i>(Conflict of interest)</i>
	Conflict of Interest to me gets patients thinking about what a conflict of interest might be. Since there is none, why even raise that question in people's minds? <i>(Conflict of interest)</i>
	Because I believe in being honest about things like that. It would come out anyway so why lie? <i>(Conflict of interest)</i>
	Telling everyone that I'm trustworthy makes me seem less trustworthy. It screams of desperation. <i>(Trust)</i>

Note: In parentheses we indicate whether the reason was for the disclosure or non-disclosure of the Physician's Conflict of Interest statement or Trust rating.

**Table 2c: Coding of reasons to disclose trustworthiness and conflict of interest statements in Experiment 6**

Disclosed	Trustworthiness	N	Strategic N (%)	Naïve Inference N (%)	Irrelevant N (%)	Ethical N (%)	Other N (%)
No	Low	22	1 (5)	18 (82)	5 (23)	0 (0)	1 (5)
	High	6	0 (0)	4 (67)	3 (30)	0 (0)	0 (0)
	<b>Total</b>	<b>28</b>	<b>1 (4)</b>	<b>22 (79)</b>	<b>8 (29)</b>	<b>0 (0)</b>	<b>1 (4)</b>
Yes	Low	4	0 (0)	1 (25)	0 (0)	3 (75)	2 (50)
	High	24	3 (13)	18 (75)	0 (0)	3 (13)	2 (8)
	<b>Total</b>	<b>28</b>	<b>3 (11)</b>	<b>19 (68)</b>	<b>0 (0)</b>	<b>6 (21)</b>	<b>4 (14)</b>
	<b>Total</b>	<b>56</b>	<b>4 (7)</b>	<b>41 (73)</b>	<b>8 (14)</b>	<b>6 (11)</b>	<b>5 (9)</b>
	<b>Kappa</b>		<b>.98</b>	<b>.80</b>	<b>.95</b>	<b>.95</b>	

Disclosed	Conflict of Interest	N	Strategic N (%)	Naïve Inference N (%)	Irrelevant N (%)	Ethical N (%)	Other N (%)
No	Strong	17	0 (0)	15 (88)	5 (29)	0 (0)	2 (12)
	None	11	0 (0)	4 (36)	8 (73)	1 (9)	1 (9)
	<b>Total</b>	<b>28</b>	<b>0 (0)</b>	<b>19 (68)</b>	<b>13 (46)</b>	<b>1 (4)</b>	<b>3 (11)</b>
Yes	Strong	7	1 (14)	0 (0)	0 (0)	6 (86)	2 (29)
	None	21	1 (5)	17 (81)	2 (10)	5 (24)	0 (0)
	<b>Total</b>	<b>28</b>	<b>2 (7)</b>	<b>17 (61)</b>	<b>2 (7)</b>	<b>11 (39)</b>	<b>2 (7)</b>
	<b>Total</b>	<b>56</b>	<b>2 (4)</b>	<b>36 (64)</b>	<b>15 (27)</b>	<b>12 (21)</b>	<b>5 (9)</b>
	<b>Kappa</b>		<b>.98</b>	<b>.77</b>	<b>.89</b>	<b>.88</b>	

Note: One response can fall into multiple categories.

**Figure 1: Stimuli in Experiment 1****INSTRUCTIONS**

In a major US city, all doctors are rated by their patients on scales from 0 -100, where 0 is extremely poor, and 100 is outstanding.

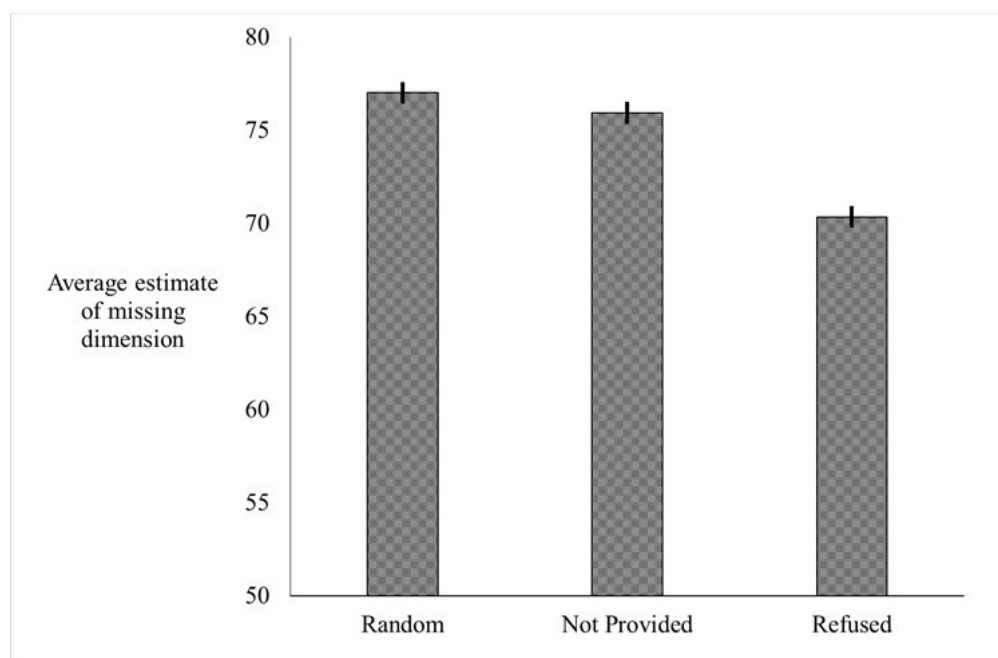
We contacted a large number of doctors and asked them to provide their average patient rating on five dimensions (quality of care, trustworthiness, availability, bedside manner and value for money).

The combined ratings for all doctors in the survey are shown here.

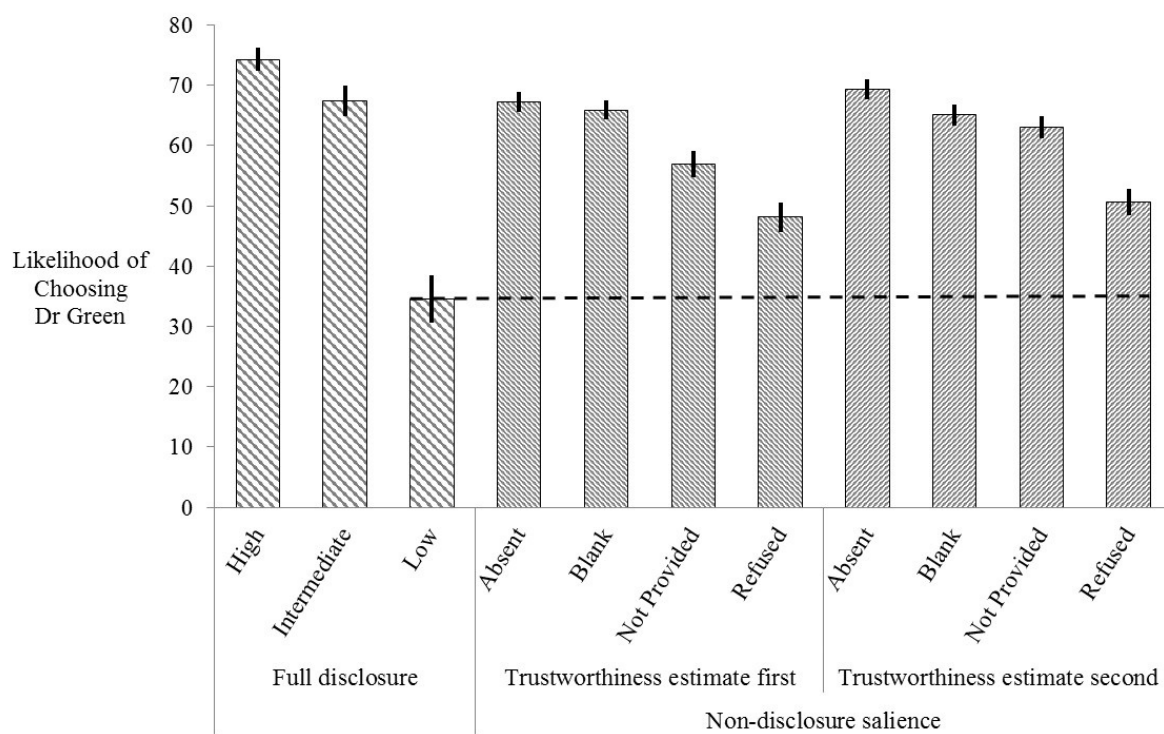
Quality of care provided	85
Trustworthiness	81
Availability	79
Bedside manner	82
Value for money	78

In this study, we will show you the average ratings for a single doctor, who we will call Doctor Y.

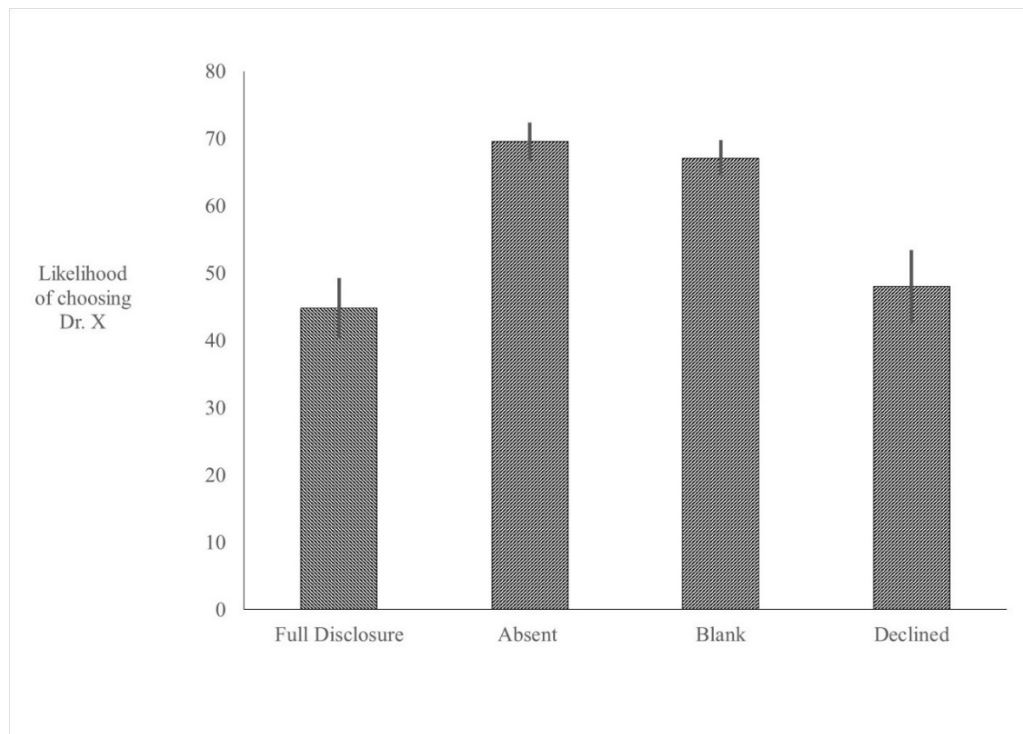
You will, however, only see four of Doctor Y's possible five ratings.

**Figure 2. Average estimate of the missing dimension in Experiment 1**

Note: Error bars are  $\pm 1$  Standard Error (SE).

**Figure 3. Likelihood of choosing Dr. Green in Experiment 2**

Note: Error bars are  $\pm 1$  SE. The dotted line indicates the unraveling prediction.

**Figure 4. Likelihood of choosing Doctor X by condition in Experiment 4**

Note: Error bars are  $\pm 1$  SE. Only the “declined” high salience non-disclosure framing is equivalent to full disclosure.